딥러닝을 활용한 CCTV 영상 기반 침수감지 기술

김길호 수자원하천연구본부 수석연구원

들어가며

최근 중국 허난(河南)성의 정저우(郑州)시에서는 기록적인 폭우로 인하여 징광터널과 지하철이 침수되어 상당한 인명피해가 발생하였다. 국내에서도 2020년 여름 부산지역에 내린 기습적인 폭우로 인하여 도심지역인 초량동 지하차도가 침수되어 안타까운 인명사고를 경험한 바 있다. 이들의 공통점은 급속한 도시화로 불투수 면적이 넓고, 도시기능적으로 지하차도 및 지하터널, 지하철, 지하통로 등과 같은 저지대 위험지역이 산발적으로 다수 존재한다는 점이다. 최근 국내에서 발생한 홍수피해 또한 하천수위 상승에 따른 외수침수보다는 저지대 내수침수로 인한 피해가 대부분이며, 내수침수는 도시지역에서 주로 발생하고 있다. 이처럼 도심에 산발적으로 위치한 침수우려지역에 대한 위험을 저감하기 위해서는 우수처리 시설용량 개선 및 확대와 같은 구조적 대책이 우선적으로 필요하며, 대응관점에서는 재난관리 담당자가 침수상황을 미리 인지하여 신속하게 대처하는 것이 필수적이다. 최근 IoT 및 4차산업 기술 발달로 인해 각종 재난을 빠르게 감지하기 위한 지능형 CCTV 활용 연구가 활발히 진행되고 있다. 이 글에서는 한국건설기술연구원에서 개발한 무인·실시간으로 운영되는 CCTV 영상 기반 침수감지 기술에 대해 소개하고자 한다.

인공신경망과 딥러닝

머신러닝 분야 중 하나인 딥러닝은 사람의 신경망 원리와 구조를 모방하여 만든 기계학습 알고리즘인 인공신경망(ANN, Artificial Neural Network)을 기초로 한다. 그림 2의 인공신경망은 다수의 입력데이터를 받는 입력층, 데이터 출력을 담당하는 출력층, 이 사이에 존재하는 은닉층이 존재하며, 역전파 알고리즘을 통해 각 노드 간의 가중치가 결정된다. 그래픽카드와 하드웨어의 발달로 ANN 고유의 문제(파라미터 최적화, 과적합, 학습시간)가 해결되면서 모델 내 은닉층을 많이 늘려서 학습의 결과를 향상시키는 심층신경망(DNN, Deep Neural Network)이 등장하였고, 통상 은닉층이 2개 이상 가지는 것을 딥러닝이라고 한다.

DNN을 응용하여 딥러닝 모델 구축에 자주 사용되는 대표적인 인공신경망은 CNN(Convolution Neural Network)과 RNN(Recurrent Neural Network)이다. RNN은 반복적이고 순차적인 데이터를 학습하는 데 특화된 인공신경망으로 구글번역, 파파고와 같은 텍스트 분류, 번역, 자연어 처리 등에 활용되며, CNN은 이미지 분류, 얼굴 인식, 텍스트 인식 등에서 주로 사용되고 있다.

본 기술에서 사용한 CNN은 데이터 특징을 추출하여 패턴을 파악하는 구조로, 이미지 또는 영상을 인식하고 분류하는 데 탁월한 성능을 가지고 있다. 그림 3의 CNN 알고리즘은 이미지의 특징을 추출하는

그림 1 CCTV 영상 기반 침수감지 기술 개요

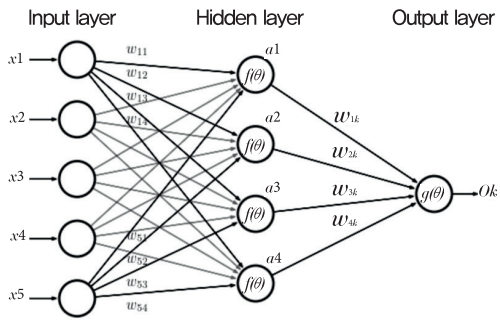


그림 2 인공신경망(ANN) 구조

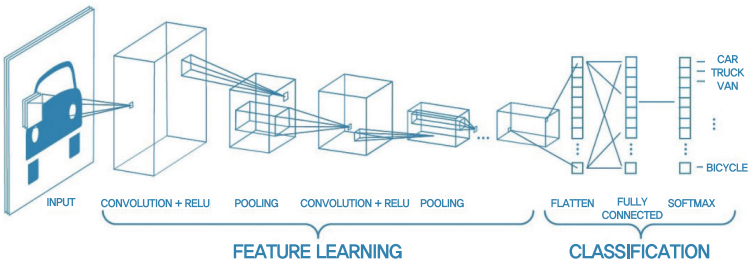


그림 3 CNN(합성곱신경망) 알고리즘의 구조

과정과 특징 및 이미지를 분류하는 과정으로 구분한다. 우선, 이미지 특징을 추출하는 과정은 이미지의 외곽선 등과 같은 특징들을 하나씩 추출하는 과정인 콘볼루션과 풀링이라는 과정을 거쳐 불필요한 요소를 제거하고 다듬어 정교한 데이터를 모델에 제공한다. 다음 분류단계는 2차원 데이터를 1차원 배열 형태로 바꿔주고 이를 이미지를 분류하는 레이어와 연결하고 최종적으로 활성함수을 통하여 이미지를 분류한다.

한편, CNN에서 다루는 객체탐지 문제 유형은 이미지 분류, 객체 탐지, 객체 세분화로 구분된다. 분류모델은 일반적으로 고정된 크기의 이미지를 입력받는 입력층을 가지며, 모델구축 방법이 간단하고 하나의 이미지에 하나의 클래스를 부여하여 학습 데이터를 쉽게 생성할 수 있다는 장점이 있다. 대표적으로 VGG-16, ResNet, ShuffleNet, SqueezeNet 등의 모델이 있다. 객체탐지 모델은 입력 이미지에서 탐지할 클래스에 해당하는 모든 물체의 개별 위치를 찾으며, 물체 위치를 박스 형태로 표시하기 때문에 물체의 위치 및 상대적 크기가 결정된다. 대표적인 모델에는 Faster R-CNN, Single Shot Detector(SSD), YOLO 등이 있다. 객체세분화 모델은 입력 이미지에서 탐지할 클래스의 물체를 모두 찾아서 픽셀단위로 세분화하기 때문에 물체의 형상을 정확하게 추출할 수 있다. 대표적인 모델에는 SegNet, DeepLab, U-Net 등이 있다.

침수판별 딥러닝 모델 개발

본 기술에서 대상으로 하는 공간은 물 객체(water body)가 항시 존재하는 하천구역이 아닌, 제내지(도로, 실내, 실외)이다. 학습데이터는 공개데이터 European Flood Dataset(Pereira 외, 2019)과 사진공유 플랫폼 Flickr, Youtube-8M으로부터 수집하였다. 수집데이터는 약 82,000개의 이미지와 600개의 영상이며, 검수 작업과 레이블 작업을 통해 딥러닝 학습 및 성능시험에 활용하였다. 학습 및 성능시험에 사용한 데이터는 도로공간 5,457개, 실외공간 2,657개, 실내공간 595개로 총 8,707개이며, 이 중 80%는 딥러닝 모형의 학습을 위한 훈련셋, 20%는 검증을 위한 시험셋으로 랜덤하게 추출하여 구분하였다.

본 기술에는 상용화 측면에서 침수판별(감지) 정확도를 담보하면서 다채널 CCTV 영상(이미지)을 실시간으로 빠르게 처리하기 위해서 Image Classification(분류모델) 모델을 적용하였다. 그리고 이미지 분류에 적합한 CNN 가운데 Kaggle에서 검증된 아키텍처인 ResNet과 ResNet50, SqueezeNet, ShuffleNet v1, ShuffleNet v2를 활용하여 침수판별 모델을 구축하였다. 일반적으로 다중 분류에서는 활성화 함수(Activation function)로 Softmax를 활용하지만, 본 기술에서는 침수의 유/무를 판별하기 때문에 시그모이드(Sigmoid)를 적용하였다. Hidden Layer 개수와 Node 개수는 각 모델들의 아키텍처를 그대로 활용

공간유형	공간정의	비고
도로공간	차량, 차선, 표지판, 신호등이 존재하는 경우 도로공간으로 표기	
실내공간	건물내부, 지하주차장 등 지붕이 존재하는 옥내를 실내공간으로 표기	
실외공간	도로와 실내에 해당하지 않은 야외의 경우 실외공간으로 표기	하천공간 제외

표 1 공간유형 구분 및 표기 정의

공간유형	구분	데이터셋(개)		
		훈련셋	시험셋	계
도로공간	Positive	2,038	510	2,548
	Negative	2,326	582	2,908
실내공간	Positive	110	28	138
	Negative	365	92	457
실외공간	Positive	911	228	1,139
	Negative	1,213	304	1,517
전체	Positive	3,059	766	3,825
	Negative	3,904	978	4,882

주) Positive: 침수, Negative: 비침수

표 2 학습데이터셋 구축 및 구분

하였고, 모델마다 최대성능을 발휘할 수 있도록 Hyper parameter를 조절하였다. 딥러닝 모델의 일반화 성능을 향상시키기 위하여, Google Research, Brain Team에서 제안한 Data Augmentation 기법을 적용하였다.

검토한 5가지 CNN 모형들은 사전에 정의된 3개의 공간유형(실외, 실내, 도로)별로 구분하여 학습하였으며, 학습과정에서 1 Epoch를 학습할 때 Batch size의 80%는 학습하고 20%에 해당하는 데이터를 검증함으로써 과적합 여부를 지속적으로 테스트하였다. 또한 과적합 방지를 위해 Early Stopping 기법을 통해 손실함수 값의 변화가 0.0001 이하로 5회 이상 진전되지 않을 경우 학습을 중단하였다. 손실함수는 로지스틱 손실인 BCEE(Binary Cross Entropy Error)를 사용하였다. 여기서 $\hat{y}_i$는 예측값, y_i는 실제값을 의미하고, 손실은 실제값이 Negative일 때 교차 엔트로피와 Positive일 때 교차 엔트로피의 합을 통해 산정할 수 있다.

$$Loss = -\frac{1}{N}\sum_{i=1}^{N} \left(y_i \times \ell n \hat{y}_i + (1-y_i) \times \ell n(1-\hat{y}_i)\right)$$

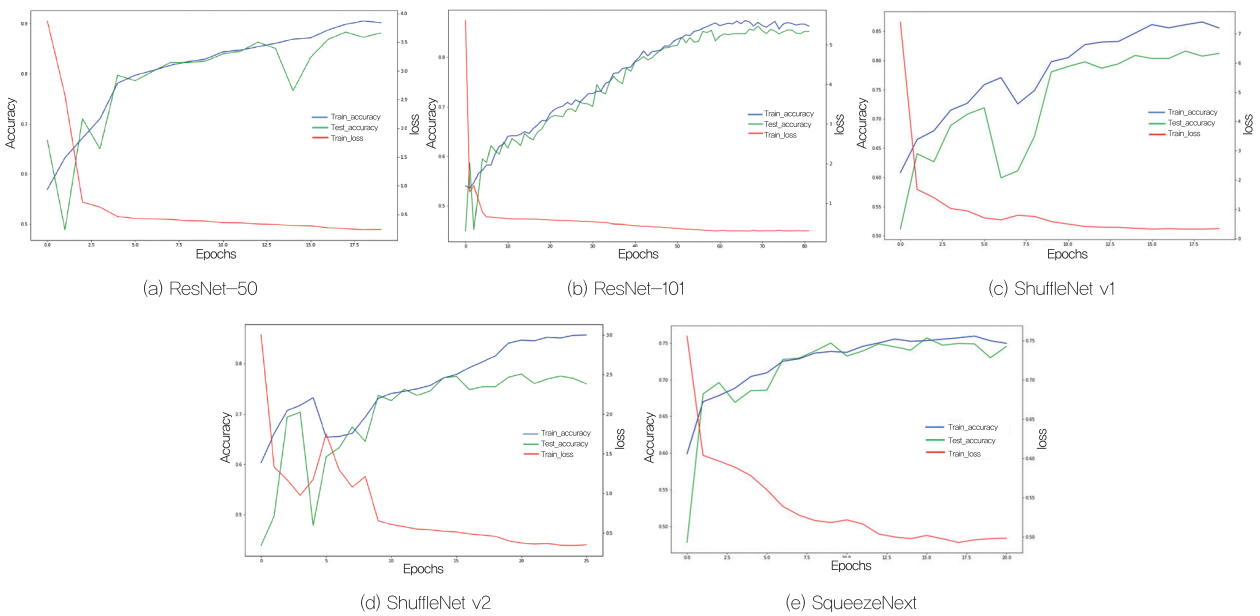
5가지 종류의 CNN 모형의 학습과정은 그림 4와 같으며, 본 연구에서 학습된 모델들의 성능을 살펴보면 Train accuracy와 Test accuracy의 추세가 유사하게 나타나는데, 대체로 과적합 문제가 발생하지 않고 적절하게 학습된 것으로 판단할 수 있다.

모델 시험평가

구축한 5가지 침수판별 딥러닝 모델을 시험셋 데이터로부터 시험평가하였다. 전체 공간에 대해서는 ResNet-50의 성능이 표 3과 같이 87.5%의 정확도로 가장 우수하게 나타났다. 실내공간의 학습데이터가 타 공간유형의 자료보다 현저하게 적기 때문에 평가결과에서도 비교적 낮은 성능을 보였다. ResNet 모형과 SqueezeNet 모형은 학습구간과 평가구간에서 약 3~4%P(% Point)차이 밖에 나지 않지만, SuffleNet_v1, SuffleNet_v2은 평가구간에서 약 10%P가량 성능이 감소하였다. 이는 SuffleNet 계열 모형들이 타 모형들에 비해 과적합 문제가 다소 있는 것으로 볼 수 있다.

추론속도 및 보급 비용 검토

수신되는 CCTV 영상을 10초당 1회 판별하는 것을 규칙으로, 모델별 경량 정도를 확인하기 위해 네트워크 파라미터 및 Batch size에 따른 모델별 추론속도를 검토하였다. 단순히 추론속도 측면에서는 1,935대까지 CCTV 채널 구성이 가능한 SqueezeNext 모델이 가장 경량한 것으로 나타났다. 침수판별 성능이 가장 우수한 것으로 확인된 ResNet-50 모델의 경우 표 4와 같이 1개 서버 당 최대 828대(채널)의 CCTV를 운용할 수 있는 것으로 나타났고, 테스트 과정에 사용한 워크스테이션 가격이 150만 원(소프트웨어 라이선스 미포함)임을 고려할 때, CCTV 1채널당 약 1,812원의 보급비용이 소요되어 상업적으로 보급하는 데 큰 문제는 없는 것으로 판단된다.



(a) ResNet-50 (b) ResNet-101 (c) ShuffleNet v1

(d) ShuffleNet v2 (e) SqueezeNext

그림 4 CNN 모델별 학습결과

구분		ResNet-50	ResNet-101	SqueezeNet	ShuffleNet_v1	ShuffleNet_v2
실외	침수 (Sensitivity)	82.0%	80.3%	70.6%	68.4%	66.7%
	비침수 (Specificity)	89.5%	86.5%	73.4%	80.9%	76.6%
	정확도 (Accuracy)	86.3%	83.8%	72.2%	75.6%	72.4%
실내	침수 (Sensitivity)	21.4%	46.4%	64.3%	46.4%	53.6%
	비침수 (Specificity)	95.7%	87.0%	63.0%	75.0%	68.5%
	정확도 (Accuracy)	78.3%	77.5%	63.3%	68.3%	65.0%
도로	침수 (Sensitivity)	87.1%	82.4%	71.8%	80.0%	73.9%
	비침수 (Specificity)	92.1%	91.8%	80.2%	86.8%	84.2%
	정확도 (Accuracy)	89.8%	87.4%	76.3%	83.6%	79.4%
공간전체(Accuracy)		87.5%	85.4%	74.6%	80.2%	76.4%

표 3 모델별 시험평가 결과

	ResNet-50	ResNet-101	SqueezeNext	ShuffleNet _v1	ShuffleNet _v2
파라미터 수	23M	43M	0.8M	3.6M	1.5M
최대 Batch size	8	8	24	16	24
추론시간 (초/Batch)	0.0966	0.1461	0.1240	0.1509	0.1309
운용채널 (대/서버)	828	548	1,935	1,060	1,833
보급비용 (원/대)*	1,812	2,737	775	1,415	818

주) 하드웨어 구축비용만 포함, 소프트웨어 라이선스와 설치비 등 미포함

표 4 모델별 추론속도 및 보급비용 검토

맺음말

이 글에서 소개한 기술은 태풍, 집중호우 등의 수재해로 인해 국지적 다발성으로 발생되는 침수상황을 CCTV 영상으로부터 자동 모니터링하고, 실시간으로 빠르게 감지하기 위한 시스템이다. 아직 현장에 적용해보지 못한 한계를 가지지만, 학습에 활용되지 않은 다양한 조건의 테스트 자료에서 모형을 평가했을 때 성능 감소가 크게 나타나지 않아 현장 적용하는 데 큰 문제가 되지 않을 것으로 예상된다. 추후 지자체와 협업을 통해 테스트베드를 결정하고, 프로토타입을 시범 운영할 계획이다. 모델 측면에서는 건물내부(실내) 학습데이터를 추가하고, 기타공간(농경지 등)으로 확대하여 보다 다양한 공간에서 활용가능할 수 있도록 고도화할 계획이다.

참고자료

- Dahal, R. K. and Hasegawa, S. (2008). Representative rainfall thresholds for landslides in the Nepal Himalaya. Geomorphology, 100(3), 429–443.
- Saito, H., Nakayama, D. Matsuyama, H. (2010). Relationship between the initiation of a shallow landslide and rainfall intensity—duration thresholds in Japan. Geomorphology, 118(1), 167–175.
- Jang, J. H. (2015). An advanced method to apply multiple rainfall thresholds for urban flood warnings. Water, 7(11)6056–6078.
- Ramos Filho, G. M., Coelho, V. H. R., da Silva Freitas, E., Xuan, Y., das Neves Almeida, C. (2021). An improved rainfall–threshold approach for robust prediction and warning of flood and flash flood hazards. Natural Hazards, Vol. 105, No. 3, pp. 2409–2429.
- Mason, D. C., Davenport, I. J., Neal, J. C., Schumann, G. J. P. and Bates, P. D. (2012). Near real-time flood detection in urban and rural areas using high-resolution synthetic aperture radar images. IEEE transactions on Geoscience and Remote Sensing, 50(8), 3041–3052.
- Baydargil, H. B., Park, J., Shin, H. S. and Park, K. (2018, October). Water flow Detection Using Deep Convolutional Encoder–decoder Architecture. In 2018 18th International Conference on Control, Automation and Systems (ICCAS) (pp. 841–843). IEEE.
- Lopez-Fuentes, L., Rossi, C. and Skinnemoen, H. (2017, December). River segmentation for flood monitoring. In 2017 IEEE International Conference on Big Data (Big Data) (pp. 3746–3749). IEEE.
- Chaudhary, P., D'Aronco, S., Moy de Vitry, M., Leitao, J. P. and Wegner, J. D. (2019). Flood-water level estimation from social media images. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 4(2/W5), 5–12.
- Pereira, J., Monteiro, J., Estima, J. and Martins, B. (2019). Assessing flood severity from georeferenced photos. In Proceedings of the 13th Workshop on Geographic Information Retrieval (pp. 1–10).