ANN 기반 도시침수 위험기준 예측 모델 개발 및 학습자료의 확장에 따른 영향 평가

Development of an ANN-Based Urban Flood Alert Criteria Prediction Model and the Impact of Training Data Augmentation

강호선* · 조재웅** · 이한승*** · 황정근**** · 문혜진*****

Kang, Hoseon*, Cho, Jaewoong**, Lee, Hanseung***, Hwang, Jeonggeun****, and Moon, Hyejin*****

Abstract

Urban flooding occurs during heavy rains of short duration, so quick and accurate warnings of the danger of inundation are required. Previous research proposed methods to estimate statistics-based urban flood alert criteria based on flood damage records and rainfall data, and developed a Neuro-Fuzzy model for predicting appropriate flood alert criteria. A variety of artificial intelligence algorithms have been applied to the prediction of the urban flood alert criteria, and their usage and predictive precision have been enhanced with the recent development of artificial intelligence. Therefore, this study predicted flood alert criteria and analyzed the effect of applying the technique to augmentation training data using the Artificial Neural Network (ANN) algorithm. The predictive performance of the ANN model was RMSE 3.39-9.80 mm, and the model performance with the extension of training data was RMSE 1.08-6.88 mm, indicating that performance was improved by 29.8-82.6%.

Key words : Urban Flood, Alert Criteria, Artificial Intelligence, Training Data Augmentation

요 지

도시침수는 단기간 높은 강도의 집중호우에 의해 단시간에 발생하기 때문에 대응을 위해서는 신속하고 정확한 침수 예·경보가 필요하다. 선행연구에서 침수피해 통계기반의 도시침수 위험기준 산정 방법을 제시하고 침수 위험기준 예측을 위한 Neuro-Fuzzy 모델을 개발하였다. 최근 인공지능 기술의 발달로 다양한 인공지능 알고리즘의 활용도와 예측 정확성이 향상됨에 따라 도시침수 위험기준 예측에 적용하여 침수 위험기준의 정확성을 향상시키고자 한다. 따라서 본 연구에서는 Artificial Neural Network (ANN) 알고리즘을 이용하여 침수 위험기준을 예측하고, 학습자료의 확장 기법 적용을 통한 영향을 분석하였다. ANN 모델의 예측 성능은 RMSE 3.39-9.80 mm, 학습자료 확장을 통해 모델 성능은 RMSE 1.08-6.88 mm로 29.8-82.6% 개선되는 것을 확인하였다.

핵심용어 : 도시침수, 위험기준, 인공지능, 학습자료 확장

*정회원, 국립재난안전연구원 책임연구원(E-mail: hydro8487@korea.kr)

Member, Principal Researcher, National Disaster Management Research Institute

**교신저자, 정회원, 국립재난안전연구원 시설연구사(Tel: +82-52-928-8175, Fax: +82-52-928-8299, E-mail: jwcho80@korea.kr)

Corresponding Author, Member, Research Officer, National Disaster Management Research Institute

***정회원, 국립재난안전연구원 책임연구원

Member, Principal Researcher, National Disaster Management Research Institute

****정회원, 국립재난안전연구원 선임연구원

Member, Senior Researcher, National Disaster Management Research Institute

*****정회원, 국립재난안전연구원 연구원

Member, Researcher, National Disaster Management Research Institute

1. 서 론

최근 기후변화로 인해 폭염, 집중호우 등 극한기상현상의 증가로 많은 피해가 발생하고 있으며, 홍수로 인한 피해도 증가하였다. 2020년에는 전세계적으로 지난 20년(2000~2019년)과 비교하였을 때 태풍은 26% 증가한 127건, 홍수는 23% 증가한 201건이 발생하였다(CRED, 2020). 특히 1월에는 인도네시아 자카르타에서 약 400 mm의 집중호우가 발생하여, 66명이 사망하고, 6만명의 이재민이 발생하였다. 또한 일본에서는 장마전선 영향으로 1,300 mm가 넘는 강우로 인해 약 80여명이 사망하였으며, 중국에서도 장마와 태풍으로 인한 4차례의 홍수로 총 397명이 사망하고 약 1,430만명의 피해가 발생하였다.

우리나라에서도 집중호우, 태풍으로 인해 해마다 많은 피해가 발생하고 있으며, 향후 기후변화에 따른 극단적 강우현상을 가속화 될 것으로 전망된다. 우리나라 연평균 강수량은 지난 100년 동안 108 mm 증가하였고, 2100년까지 그 2배인 216 mm가 더 증가할 것으로 예측하고 있다. 또한 강한 강수는 증가하고 약한 강우는 감소하는 현상으로 극단적인 강우를 나타내는 지수인 5일 최대강우량, 강우강도의 증가가 예상된다(Korea Meteorological Administration, 2021). 2020년에는 1973년 이후 가장 긴 장마(중부 기준 54일)와 함께 8~9월 연이은 4개 태풍의 영향으로 1조 3,585억원의 재산피해와 46명의 인명피해가 발생하였으며, 이는 최근 10년(2010~2019) 연평균 피해(재산 3,883억원, 인명 14명) 대비 약 3배를 초과하는 수치이다(Korea Meteorological Administration, 2021).

행정안전부 국가재난관리정보시스템(Natural Disaster Management System, NDMS)에서는 2020년 태풍, 호우 등 16건 재해로 총 3,223건의 침수피해가 발생하였다. 특히, 도시지역에서는 저지대, 지하차도, 우수관 역류, 배수펌프장 미작동 등으로 인해 침수가 발생하였다. 이는 도시침수가 지속시간 1~2시간 이내 단시간 집중호우에 의해 발생하여 신속한 대응 및 정확한 예측이 어렵기 때문이다.

도시침수 예측을 위해 서울시에서는 침수예측시스템을 개발하여 2차원 침수해석을 통해 강우량과 지속시간에 따라 80개의 시나리오를 작성하여 침수를 예측하고 있으며, 한국과학기술정보연구원(KISTI)에서는 인천시 전체 배수분구를 대상으로 2차원 침수해석을 통해 20여종의 시나리오를 생산하여 침수예측에 활용하고 있다. ‘SWMM 이중배수모형(1D-1D) 기반 실시간 도시침수 예측(Jang et al., 2020)’에서는 실시간 도시침수 예측을 위해 빠른 예측이 가능한 SWMM 1D-1D 모형을 사용하였으며, ‘레이더 예측강우를 활용한 도시침수예측(Yoon et al., 2014)’에서는 예측강우와 SWMM 모형을 사용하였다. ‘LSTM 모형과 로지스틱 회귀를 통한 도시 침수 범위의 예측(Kim et al., 2020)’에서는 SWMM 유출 모의 결과를 활용하여, 해당구역의 침수범위를

예측하였다. 이와 같이 최근 도시침수 예측을 위해 등 많은 연구가 이루어지고 있으나, 대부분 SWMM 등 프로그램을 이용한 관망해석, 지표유출 등 수치해석을 통한 연구가 대부분이다. 그러나 이와 같은 방법은 전국을 대상으로 적용하기에는 한계가 있으며, 빠르게 변화하는 도시구역의 특성을 수시로 반영하는 것 또한 매우 어려운 작업이다.

선행연구 ‘지역별 주요 재난대응 시나리오 및 기술 개발’에서는 침수피해이력과 강우자료를 이용하여 통계기반의 도시침수 위험기준 산정 방법을 제시하였으며, 유역특성자료와 침수 위험기준과의 관계를 통해 Neuro-Fuzzy 알고리즘을 이용한 예측 모델을 개발하였다(National Disaster Management Research Institute, 2014; 2015; 2016; 2017). 또한 초기화 함수 및 학습자료 구성에 따른 영향을 분석하고 모델을 개선하였다(Kang et al., 2020).

그러나 최근 인공지능 기술의 발달로 다양한 딥러닝 및 머신러닝 알고리즘의 활용도와 예측 정확성이 향상됨에 따라 도시침수 위험기준 예측에 적용하고자 한다. 따라서 본 연구에서는 Artificial Neural Network (ANN) 알고리즘을 이용하여 도시침수 위험기준 예측하고, 학습자료의 확장 기법 적용을 통한 영향을 분석하여 예측 정확성을 향상시키고자 한다.

2. 연구내용 및 방법

도시침수 위험기준 예측을 위해 Artificial Neural Network (ANN) 알고리즘을 사용하였다. ANN은 인간의 뇌를 모방하여 인간의 사고방식을 컴퓨터에 학습시키는 알고리즘으로 뇌의 기본단위인 뉴런은 연결된 뉴런들로부터 정보를 받아 고유한 방식으로 처리하고, 그 결과를 다른 뉴런에 전달하는 뇌의 거대한 생물학적 네트워크 구조를 모방하여 주어진 문제를 해결하기 위해 만들어진 알고리즘이다. 뉴런의 기능을 컴퓨터에서 재현하기 위해 Fig. 1과 같이 첫 번째 층인 입력층(Input layer)으로 들어온 n 개의 데이터를 각 뉴런에서 특정 가중치를 곱한 후 합하여 다음 뉴런과 은닉층(Hidden layer)로 이동시키며, 마지막 층인 출력층(Output layer)의 뉴런에서 결과를 출력한다.

본 연구에서는 입력변수로 불투수율(Impermeability rate), 유역경사(Basin slope), 관거밀도(Culvert density), 빗물받이 밀도(Inlet density)를 사용하였으며, 결과변수로 과거 침수 피해이력기반 위험기준 산정 결과값을 사용하여 총 315개의 Training Data Set A를 구축하였다.

입력변수는 선행연구에서 도시유출 특성에 영향을 주는 변수들에 대한 상관분석 및 민감도 분석을 통해 선정된 유역특성을 사용하였다. 불투수율은 환경부 환경공간정보 서비스에서 제공하는 2018년 토지피복도를 이용하여 산정하였으며, 유역경사는 국토부 국토정보플랫폼에서 제공하는 2014년 Digital Elevation Model (DEM), 관거밀도와 빗물

Fig. 1. Artificial Neural Network

Table 1. Training Data Set A

District	Basin Characteristics Parameter				Flooding Limit Rainfall by Duration (mm)		
	Impermeable Rate (%)	Basin Slope (%)	Culvert Density	Inlet Density	30 min	60 min	180 min
Ulsan Jung-gu Jungang-dong	62.55	4.12	0.001631	0.001085	38.20	54.90	97.70
Busan Yeonje-gu Geojel-dong	88.20	1.45	0.027635	0.000395	37.63	51.84	86.11
Gyeonggi-do Gwangmyeong-si Gwangmyeong2-dong	45.541	9.87	0.000632	0.000106	38.28	54.91	97.26
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Minimum	1.31	0.00	0.000003	0.0000001	23.31	30.71	42.78
Maximum	100.00	48.96	0.052182	0.0055238	66.92	94.60	214.47
average	52.02	11.58	0.009456	0.0006370	43.02	61.15	109.11

받이 밀도는 각 지자체별로 수집한 지하시설물 자료를 이용하였으며, GIS 분석은 QGIS를 사용하여 분석하였다.

결과변수인 과거 침수피해이력기반 위험기준은 선행연구(National Disaster Management Research Institute, 2014; 2015; 2016; 2017)에서 제시한 피해이력기반 한계강우량 산정 방법을 이용하여 산정하였으며, 2013~2019년 동안의 침수피해이력과 강우자료를 사용하였다. 여기서, 침수 위험기준은 침수가 발생(약 20 cm 정도)했을 때의 강우량으로 정의하였다. 침수피해자료는 행정안전부 국가재난관리시스템에 등록된 자료를 사용하였으며, 강우자료는 기상청에서 제공하는 ASOS, AWS, AMOS 545개 관측소의 1분 단위 강우 관측데이터를 수집하여 10분, 30분, 60분, 180분, 360분,

720분 지속시간에 대한 최대누적강우량을 분석하였다.

Training Data Set A의 일부를 Table 1에 나타내었으며, 입력변수의 범위는 불투수율 1.31~100.0 유역경사 0.0~48.96, 관거밀도 0.000003~0.052182, 빗물받이 밀도 0.0000001~0.0055238이다.

Training Data Set A를 ANN 알고리즘에 적용하여 Model A를 개발하였으며, 교차검증을 통해 모델을 평가하였다. 다음으로 학습자료 확장을 위해 Fig. 2와 같이 Model A의 예측 결과를 검증하여 Training Data Set B를 구축하고 이를 활용해 Model B를 개발하였으며, 마지막으로 Model A와 B의 예측성능을 비교하여 학습자료 확장에 따른 영향을 분석하였다.

Fig. 2. Research Process

Table 2. ANN Algorithm Parameter

Parameter	Maximum	Minimum	Cases
Learning rate	0.01	0.00001	5
alpha	0.1	0.000001	5
tolerance	0.01	0.00001	5
epsilon	0.000001	0.0000000001	5
dropout rate	0.0	0.1	-
kernel_ini, bias_ini	uniform, zero, Constant, Normal, one		

예측 성능 평가를 위한 검증방법으로는 K-fold 교차검증 방법을 이용하였다. K-fold는 전체 데이터를 k개의 집합으로 나누어 k-1개의 집합을 학습자료로 사용하고 나머지 1개의 집합을 검증데이터로 사용하는 방법을 순차적으로 반복하여 평가하는 방법으로, 모든 데이터가 학습 및 검증자료로 사용되며 과적합이 일어날 확률이 낮다. 본 연구에서는 fold의 개수를 10개로 하여 각각의 모델에 대한 예측 결과를 통합하여 제시하였으며, 모델 성능은 Eqs. (1)~(3)를 이용하여 평균절대오차(Mean Absolute Error, MAE), 평균제곱근오차(Root Mean Square Error, RMSE), 평균절대비율오차(Mean Absolute Percentage Error, MAPE)를 비교하였다.

여기서, y 는 실제값, $\hat{y}$ 는 예측값, $\bar{y}$ 는 실제값의 평균, $\bar{\hat{y}}$ 는 예측값의 평균, n 은 위험기준 개수이다.

$$MAE = \frac{1}{n} \sum |y - \hat{y}| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum (y - \hat{y})^2} \quad (2)$$

$$MAPE = \frac{1}{n} \sum \left| \frac{y - \hat{y}}{y} \right| \quad (3)$$

3. 도시침수 위험기준 예측 모델 개발

ANN 알고리즘을 사용하기 위해 Tensorflow와 Keras를

적용하였다. Tensorflow는 Google에서 개발한 것으로 이식성에 높은 인공지능망 라이브러리이다. 특히, 딥러닝에서 많이 사용하는 다양한 모델과 알고리즘과 Tensorboard가 포함되어 있어 분석결과를 Data Flow Graph (DFG)로 표현할 수 있다. Keras는 파이썬을 이용하여 작성된 신경망 기반의 오픈소스 라이브러리이며, Deeplearning4j, Tensorflow, Microsoft Cognitive Toolkit 및 Theano에서 사용할 수 있다. Keras는 신경망과 빠른 계산이 가능하도록 설계되었으며, 모듈 방식으로 확장성이 높다. 또한, Tensorflow에서 작동하는 것으로 파라미터 설정이나 신경망 구조를 간단하게 설계할 수 있다.

알고리즘 매개변수인 dropout rate, kernel 및 bias의 초기값 등을 Table 2와 같이 설정하였다. dropout 비율은 모델의 과적합을 방지하기 위해 학습시 임의의 비율만큼 뉴런을 제외하여 학습시키기 위한 변수이며, Kernel 및 bias의 초기값은 그라디언트 소멸과 폭발문제를 해결하기 위해서이다. 위 세가지 매개변수를 변경해가며 최적의 지속시간별 한계강우량을 추정하였다. ANN을 이용하여 한계강우량을 추정하기 위한 활성화 함수로 Rectified Linear Unit (ReLU)를 사용하였다. 활성화 함수는(Activation Function)는 노드에 입력된 값들을 비선형 함수에 통과시킨 후 다음 레이어로 전달하는데 사용되는 함수를 의미한다(Renew's blog CS and deep learning, 2018). ReLU 외에 Sigmoid, Tanh 등 다양한 활성화 함수가 있지만 대부분의 함수들이 0.0~1.0 사이값을 출력하며, 학습이 진행되면서 1.0 이하의 값이 반복하여 곱해지면 0에 수렴할 가능성이 있기 때문에 ReLU

를 사용하였다. ReLU는 입력값이 0 이상이면 입력된 값을 출력하고 0 이하의 값은 0으로 반환한다(Saito, 2019). 손실함수는 가능한 작은 값을 가질 수 있도록 만드는 최적화 함수는 Momentum과 AdaGrad (Adaptive Gradient)가 혼합된 방법인 Adam 함수를 사용하였다. Momentum은 경사하강을 통해 최적값으로 이동하는 과정을 더욱 빠르게 해주며, AdaGrad는 변수들을 갱신할 때 변수마다 Step size를 다르게 설정해서 최적값으로 이동하는 방식이다. 최댓값과 최솟값의 범위를 10의 거듭제곱으로 적용하여, 뉴런의 개수는 층별 5개, 10개, 은닉층의 수는 2-3층, 매개변수별 조건은 625개로 총 2,500개의 학습조건을 설정하였다.

ANN 알고리즘에 구축된 Training Data Set A를 적용하여 도시침수 위험기준 예측을 위한 Model A를 개발하였으며, 모델 성능평가 결과는 Table 3과 같이 지속시간별로 평균절대오차는 2.97~7.05 mm, 평균제곱근오차는 3.39~9.80 mm, 평균절대비율오차는 7.07~8.31%로 분석되었다.

Table 3. Model A Prediction Performance

Parameter	Rainfall Duration		
	30 min	60 min	180 min
MAE (mm)	2.97	4.99	7.05
RMSE (mm)	3.39	5.95	9.80
MAPE (%)	7.07	8.31	6.80

Table 4. Verification Method of Data

Division	Verification Method
True (T)	$D_{min} > P > R_{max}$
Available (A)	$P > D_{min} > R_{max}$
False (F)	$P < R_{max}$

Table 5. Verification Result

Division		Number of Data	Verification Result
True, True, True		48	Application
True, True, Available		44	
True, True, False		82	
True, Available, Available		28	
Available, Available, Available		289	
Available, Available, False	True	327	Application of Reconsideration
	False	21	Reject of Reconsideration
True, Available, False		38	Reject
True, False, False		141	
Available, False, False		481	
False, False, False		619	

4. 학습자료 확장에 따른 영향 분석

4.1 학습자료의 확장

Model A를 통해 예측된 결과를 검증하여 다시 학습자료로 사용하는 방법을 적용하였으며, 확장된 학습자료를 사용하여 개발된 Model B의 성능을 평가하였다. 먼저 Model A를 통해 유역특성자료가 구축된 2,118개 행정동에 대한 도시침수 위험기준을 예측하였으며, 과거 침수피해와 강우자료를 이용하여 검 · 보정하여 확장된 Training Data Set B를 구축하였다.

Table 4의 조건을 이용하여 모델 결과를 검증하였다. P 는 예측된 위험기준, R_{max} 는 해당지역의 피해가 발생하지 않은 최대강우량, D_{min} 은 침수피해가 발생했을 때 최소강우량을 의미한다. 예측된 위험기준이 R_{max} 보다 크고, D_{min} 보다 작을 경우 True, R_{max} 와 D_{min} 보다 큰 경우 Available, R_{max} 보다 작은 경우 False로 판단하였다.

위 조건에 따라서 지속시간 30분, 60분, 180분에 대한 유효성을 판단하였으며, 3개의 값 중 2개 이상이 True인 경우, True가 1개, Available가 2개 또는 Available가 3개인 경우에는 학습자료로 사용하였다. 또한 Available 2개 False가 1개인 경우에는 P 값을 R_{max} 로 수정하여 동일한 방법으로 재검토하였으며, 이외 조건은 데이터의 신뢰성을 확보할 수 없어 학습자료에서 제외하였다. 학습자료 적용 가능 여부를 검토한 결과 Table 5와 같이 2,118개 중 818개 데이터가 학습자료로 사용 가능 한 것으로 분석되었다.

Training Data Set A와 검증된 818개 데이터의 각 변수를 지속시간별로 비교하여 Figs. 3~5에 나타내었다. Training Data Set A와 비교하여 유역특성과 한계강우량간의 상관성이 명확해지는 것을 확인할 수 있으며, 유역특성과 한계강우량의 관계에서 양의 상관관계를 보이고 있지만 지속시간 180분에서는 관거밀도와 불투수면적은 음의 상관관계를 가지는 것을 확인할 수 있었다. 이는 기존 Training Data

Set A와 유사한 경향을 보이고 있어 검증된 데이터는 학습자료로 활용 가능한 것으로 판단된다.

따라서, 기존 Training Data Set A와 검증된 Data 818개를 추가한 Training Data Set B를 구축하였으며, Table 6에 최소, 최대, 평균값을 제시하였다. 입력변수의 범위는 불투수율 0.04~100.0, 유역경사 0~47.58, 관거밀도 0.0000003~0.0575376, 빗물받이 밀도 0~0.0078773이다.

4.2 도시침수 위험기준 예측 성능 및 영향 평가

Training Data Set B를 활용하여 도시침수 위험기준을 예측한 결과를 Table 7과 Fig. 6에 제시하였다. 지속시간 30분에 대한 예측결과는 평균절대오차 1.06 mm, 평균제곱근오차 2.00 mm, 평균절대비율오차 2.42%로 나타나 예측 성능이 매우 우수한 것을 확인할 수 있었지만 한계강우량이 높은 구간에서 과대 추정되는 범위가 있으며, 전체 구간에서 예측값이 실측값보다 낮게 추정된 것을 확인할 수 있었다.

Fig. 3. Comparison of Training Data Set A and B (Duration 30 min)

Fig. 4. Comparison of Training Data Set A and B (Duration 60 min)

Fig. 5. Comparison of Training Data Set A and B (Duration 180 min)

Table 6. Training Data Set B

District	Basin Characteristics Parameter				Flooding Limit Rainfall by Duration (mm)		
	Impermeable Rate (%)	Basin Slope (%)	Culvert Density	Inlet Density	30 min	60 min	180 min
Minimum	0.04	0	0.0000003	0	35.00	49.68	75.71
Maximum	100.00	47.58	0.0575376	0.0078773	53.70	83.56	129.30
average	50.09	11.76	0.0069889	0.0009614	41.59	57.04	96.06

지속시간 60분 예측결과는 평균절대오차 0.41 mm, 평균제곱근오차 1.03 mm, 평균절대비율오차 0.71%로 나타나 지속시간 30분 예측성능보다 우수하였으며, 일부 예측결과가 실제값보다 낮은 것을 제외하고는 매우 유사한 것을 확인할 수 있다.

지속시간 180분 예측결과는 평균절대오차 4.29 mm, 평균제곱근오차 6.88 mm, 평균절대비율오차 4.37%로 분석되어 지속시간 30분, 60분 결과에 비해 예측성능이 낮은 것으로 분석되었다.

Model A와 Model B의 예측 성능을 비교하여 학습자료 확장에 따른 영향을 분석하였으며, 그 결과를 Table 8에

Table 7. Model B Prediction Performance

Parameter	Rainfall Duration		
	30 min	60 min	180 min
MAE (mm)	1.06	0.41	4.29
RMSE (mm)	2.00	1.03	6.88
MAPE (%)	2.42	0.71	4.37

나타내었다. MAE는 지속시간 30분에서 1.90 mm, 60분 4.58 mm, 180분 2.76 mm 감소하였으며, RMSE는 각각 1.39 mm, 4.91 mm, 2.92 mm, MAPE는 4.65%, 7.60%, 2.43% 감소하여 모델 성능이 향상되는 것으로 분석 되었다. 지속시간 60분 예측결과가 예측 성능도 가장 좋은 것으로 나타났으며, 개선효과도 가장 높은 것으로 나타났으며, 지속시간 180분에서 개선효과가 가장 낮은 것으로 나타났다.

마지막으로 Model A, B와 선행연구에서 개발된 Neuro-Fuzzy 모델의 예측성능을 비교하기 위해 지속시간 60분과 180분에 대한 평균제곱근오차를 비교하여 Table 9에 나타내었다. Neuro-Fuzzy 예측 성능과 비교한 결과에서도 평균제곱근오차를 기준으로 Neuro-Fuzzy Model 대비 Model A는 지속시간 60분에 42.2%, 180분에 46.2%, Model B는 90.0%, 62.1% 개선되는 것으로 나타났다.

5. 결 론

본 연구에서는 강우기반의 침수위험기준을 설정하기 위

Fig. 6. Model B Prediction Performance Graph

Table 8. Improvement Effect by Training Data Augmentation

Division	Model A - Model B			Decreasing Rate (%)		
	30 min	60 min	180 min	30 min	60 min	180 min
MAE (mm)	-1.90	-4.58	-2.76	64.2	91.8	39.1
RMSE (mm)	-1.39	-4.91	-2.92	41.0	82.6	29.8
MAPE (%)	-4.65	-7.60	-2.43	65.8	94.5	35.7

Table 9. Comparison of Model A, B and Neuro-Fuzzy Model

Model	Number of Training Data	RMSE (mm)		Decreasing Rate (%)	
		60 min	180 min	60 min	180 min
Neuro-Fuzzy Model (Kang et al., 2020)	118	10.3	18.2	-	-
Model A	315	5.95	9.80	42.2	46.2
Model B	1,133	1.03	6.90	90.0	62.1

해 비선형적인 관계를 가지는 유역특성을 이용하여 ANN 알고리즘을 이용한 예측 모델을 개발하였다. 또한 Training Data가 부족한 한계점을 개선하기 위해 데이터 확장기법을 통해 모델의 예측성능을 개선하였다.

불투수율, 유역경사, 관거밀도, 빗물받이 밀도를 입력변수로, 과거 침수피해이력기반 위험기준 산정 결과를 결과변수로 사용하여 모델 개발을 위한 Training Data Set A를 구축하였으며, ANN 알고리즘에 적용하여 도시침수 위험기준 예측을 위한 Model A를 개발하였다. K-fold 교차검증을 통해 모델 성능을 평가한 결과 지속시간에 따라 평균절대오차 2.97~7.05 mm, 평균제곱근오차 3.39~9.80 mm, 평균절대비율오차 7.07~8.31%로 분석되었다.

학습자료 확장을 위해 Model A를 통해 예측된 2,118개의 도시침수 위험기준을 과거 침수피해와 강우자료를 활용하여 검 · 보정하여 818개의 학습자료를 추가로 구축하였다.

Training Data Set A에 확장된 데이터를 추가하여 Training Data Set B를 구축하였으며, 총 1,133개 학습자료를 활용하여 Model B를 개발하였다. 개선된 모델의 예측 성능은 평균절대오차 0.41~4.29 mm, 평균제곱근오차 1.03~6.88 mm, 평균절대비율오차 0.71~4.37%로 나타나, Model A 대비 지속시간 30분에서 41.0~65.8%, 60분 82.6~94.5%, 180분 29.8~39.1% 개선되는 것으로 나타났다. 또한 선행연구인 Neuro-Fuzzy 모델과 비교한 결과에서도 평균제곱근오차 기준 지속시간 60분 90.0%, 180분 62.1% 개선되었다.

ANN 알고리즘과 Training Data 확장을 통한 도시침수 위험기준 예측 모델을 개선 효과를 확인하였으며, 향후 개선된 Model B를 활용하여 도시침수 위험기준을 예측한다면 보다 정확하고 효율적인 재난상황관리가 가능할 것으로 판단된다.

본 연구를 통해 기본적으로 방대한 데이터를 필요로 하는 딥러닝 알고리즘에 비해 Training Data가 부족한 한계점을 극복하기 위해 데이터 확장 기법을 사용하였으며, 이를 통해 모델을 성능이 향상되는 것을 확인 할 수 있었다. 선행연구에서와 같이 Training Data가 증가하면 모델의 성능이 향상되는 동일한 결과를 도출하였으며, 향후 데이터 추가 구축을 통해서도 지속적으로 개선이 가능할 것으로 판단된다. 또한 매년 발생 침수피해자료와 강우자료를 활용하여 예측된 침수위험기준을 검증 및 개선 할 필요가 있으며, 특히 정확한 침수시간을 파악할 수 있는 침수센서, CCTV 영상 등 침수자료를 활용한다면 침수 위험기준의 정확성이 향상될 수 있을 것이다.

References

Centre for Research on the Epidemiology of Disasters (CRED). (2020). *Disaster Year in Review 2020*.

- Jang, J.K., Park, M.K., Lee, N.E., Lee, J.M., and Yang, D.M. (2020). Real-time prediction of urban inundation based on SWMM 1D-1D model. *Journal of the Korea Society of Hazard Mitigation*, Vol. 2020, No. 103, pp. 401-411.
- Kang, H.S., Cho, J.W., Lee, H.S., and Hwang, J.G. (2020). Improvement of urban flood alert criteria prediction model based on neuro-fuzzy initial function and training data. *Journal of the Korea Society of Hazard Mitigation*, Vol. 20, No. 1, pp. 327-337.
- Kim, H.I., Han, K.Y., and Lee, J.Y. (2020). Prediction of urban flood extent by LSTM model and logistic regression. *Journal of Civil and Environmental Engineering Research (KSCE)*, Vol. 40, No. 3, pp. 273-283.
- Korea Meteorological Administration. (2021). *Abnormal climate report 2020*.
- National Disaster Management Research Institute. (2014). *A development of regional major disasters response scenarios and standards (I): Focused on urban flooding and drought*.
- National Disaster Management Research Institute. (2015). *Establishment of foundation for regional urban flood response system (I)*.
- National Disaster Management Research Institute. (2016). *Establishment of foundation for regional urban flood response system (II)*.
- National Disaster Management Research Institute. (2017). *Development of technology on natural disaster risk assessment and urban flood forecasting*.
- Reniew's blog CS and deep learning. (Jul 1, 2018). *Activation function used in deep learning*. Retrieved Nov 30, 2017, from <https://reniew.github.io/12/>
- Saito, G. (2019). *Deep learning from scratch*. Seoul: Hanbit Media. Inc.
- Yoon, S.S., Bae, D.H., and Choi, Y.J. (2014). Urban inundation forecasting using predicted radar rainfall: Case study. *Journal of the Korea Society of Hazard Mitigation*, Vol. 14, No. 3, pp. 117-126.

Received	October 19, 2021
Revised	October 19, 2021
Accepted	October 27, 2021