

시간당 특성 변화량 분석 신경망을 통한 태풍 예측 모델

강은솔¹, 박준혁^{2,3}, 백지오², 최윤서¹

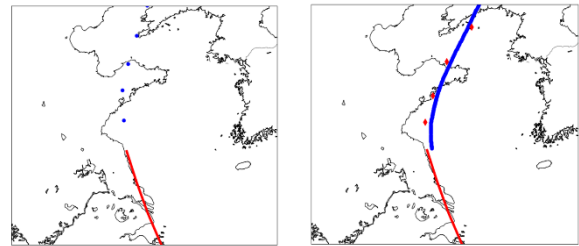
¹소프트웨어학과 ²컴퓨터공학과 ³바이오융합공학전공

초록

한반도를 비롯한 동아시아 지역에서 태풍은 매년 막대한 인명 및 경제적 피해를 야기하는 중대한 자연 재해이다. 이러한 태풍을 예측하고 대비하기 위하여 각 국가의 기상 예보 기관은 고성능의 슈퍼 컴퓨터와 복잡한 예보 모델을 활용한 예측을 수행하고 있으나, 이러한 방법은 고비용의 컴퓨팅 자원과 대규모의 기상 데이터를 요구하며 오랜 시간의 연산이 수행되는 한계가 존재한다.

이러한 한계를 극복하기 위해 최근 몇 년간 놀라운 발전을 보인 딥러닝 기술을 태풍이나 산불과 같은 자연 재해 예측에 활용하는 연구가 국내외에서 시도되고 있다. 특히, 딥러닝 기반 태풍 예측의 경우 국내에서 다양한 기관에 의한 연구가 진행되었으나, 예측의 형태가 학습에 사용된 데이터의 형태에 의하여 제약되는 딥러닝 방법의 특성 상 각 모델은 원하는 만큼의 미래가 아닌, 6시간, 12시간, 24시간과 같이 6시간 간격의 미래에 대한 예측만을 수행할 수 있으며, 학습과 평가를 위한 명확한 기준이나 데이터가 부재하여 각 연구가 별도의 기준과 방법으로 진행되고 있는 실정이다.

본 연구에서는 태풍의 과거와 현재 데이터 간의 시간적 특성 변화량을 활용하여 원하는 만큼의 시간이 흐른 후의 태풍 상태를 예측하는 심층 신경망 구조인 시간당 특성 변화량 분석 신경망(Temporal Feature Difference Analysis Network, TFDAN)을 제안한다. TFDAN 구조의 신경망은 서로 다른 두 시점에 측정된 태풍의 데



〈그림 1. 기존 태풍 예측 모델(좌)과 TFDAN(우)의 예측 형태 비교, 실제 태풍의 이동 경로는 붉은 색 마름모로 표시하였다.〉

이터로부터 얻어진 태풍의 특성 간의 차이를 이용하여 태풍의 시간당 특성 변화량을 얻고, 이를 활용하여 원하는 만큼의 미래의 태풍의 상태를 예측한다. 이러한 방법을 통해 일반적으로 6시간이나 3시간 단위로 존재하는 태풍 데이터의 형태와 무관하게 자유로운 시간(1시간, 10분 등) 간격으로 미래의 태풍을 예측할 수 있다.

우리는 이 논문을 통해 앞서 소개한 데이터의 시간적 구성에 제약 받지 않는 시계열 분석 신경망 구조와 태풍 관련 연구를 위한 데이터셋을 제안한다. 또한, 우리는 제안하는 모델을 활용하여 기상 지식이 적은 비전문가 일반인들도 어렵지 않게 태풍 예측 정보를 확인할 수 있는 서비스인 “게셋거라”를 개발하였다.

1. 서론

태풍은 주로 적도 부근에서 열대성 저기압의 발달로 발생하며, 지역에 따라 사이클론이나 허리케인 등으로 부르기도 한다. 우리나라를 비롯하여 동아시아 지역에 영향을 주는 태풍은 일반적으로 동경 180도 인근의 남중국해에서 발생하며, 7월에서 10월 사이에 가장 흔히 발생한다.[1]

이러한 태풍은 매년 한반도를 비롯한 동아시아 지역에 막대한 인적·재산적 피해를 야기하기에 태풍의 향후 진로와 세기를 정확하게 예측하여 민관 차원에서 대비할 수 있도록 하는 기술 및 서비스가 요구된다.

한편, 최근 놀라운 발전을 이루고 있는 딥러닝 기술을 태풍 예측에 활용하여, 기존 방법론들에 비해 저비용, 고성능으로 태풍을 예측해보고자 하는 연구가 진행되고 있으나, 대규모의 데이터를 이용한 지도학습을 통해 진행되는 딥러닝 기술 특성상 두 가지 한계에 직면한 상황이다.

하나는 딥러닝 기반 태풍 예측 연구를 위한 표준 데이터(baseline)의 부재이다. 태풍 예측에 활용할 수 있는 데이터는 기후 데이터, 위성 영상 데이터 등 다양하나 아직까지 태풍 예측 과제에 적합한 표준화된 데이터와 평가 지표가 존재하지 않아, 각 연구가 서로 다른 데이터와 지표로 평가되었다. 이로 인해 각 연구 결과를 비교하고 기존 연구에 이은 후속 연구가 진행되기 어려운 상황이었다. 다른 문제는 대표적인 태풍 데이터인 **best track** 데이터의 형식에 따른 모델의 예측 범위 제약 문제이다. 태풍의 중심 위치를 비롯한 정보가 담긴 **best track** 데이터는 한국 기상청, 일본 기상청, 미 합동 태풍 정보 센터 등 다양한 출처에서 제공되는데, 대부분 6시간이나 3시간 간격으로 측정된 데이터로 구성되어 있다. 이에 따라 **best track** 데이터로 학습된 태풍 모델들은 6시간 단위로 태풍의 이동경로를 예측하게 되는 한계가 있었다.

본 연구에서는 위에서 언급한 문제들에 대한 해결책으로 먼저 태풍 연구를 위한 공개 데이터셋과 베이스라인을 제공하고, 데이터의 시간적 구성과 무관하게 자유로운 시간 후의 태풍을 예측하는 모델 구조를 제안한다. 먼저, 공개 데이터셋은 일본 기상청의 **best track** 데이터를 기계학습 분야에서 흔히 사용하는 **csv** 형태로 가공하여 공개 데이터셋 공유 서비스인 **kaggle**을 통해 배포하였으며, 성능 분석을 위한 심층 신경망 기반 베이스라인을 개발하여 함께 제공하였다. 두 번째 문제를 극복하기 위한 방법으로는, 학습 데이

터의 시간적 구성과 관계없이 원하는 시간이 지난 후의 태풍을 예측하기 위한 심층신경망 구조인 시간적 특성 변화량 분석 신경망 구조를 제안한다. 이 구조의 모델은 현재와 과거 두 시점의 태풍으로부터 특성 벡터를 추출하여 두 벡터의 차이를 단위 시간으로 나누어 시간당 특성 변화량을 계산한다. 이 변화량을 원하는 시간만큼 곱하여 예측에 활용하는 구조를 통해 6시간 단위로 구성된 **best track** 데이터로 학습된 모델임에도, 원하는 만큼의 시간(1시간 30분, 25분 등) 후의 태풍 경로를 예측할 수 있는 모델을 개발하였다. <그림 1>의 좌측에 표기된 기존 모델의 예측은 6시간 단위로 단절된 예측을 수행하는 반면, <그림 1> 우측의 TFDAN은 1분 단위로 예측을 수행하여 태풍의 상세한 경로를 예측함을 확인할 수 있다. 모델의 구조에 대한 자세한 내용은 본 논문의 4절에서 소개하였다.

제안된 모델은 6시간 전과 현재의 태풍 데이터를 기반으로 태풍의 미래 진로를 예측하여, 기존 딥러닝 회귀 모델 기반 태풍 예측 연구[3]의 6시간 후 태풍 경로 오차 176km 대비 약 68% 뛰어난 55.1km의 오차로 태풍을 예측하였다. 태풍의 중심부가 대략 30km 정도의 반경을 가지는 것을 고려하면, 예측 결과가 상당히 정확함을 알 수 있다.

더불어, 우리는 제안된 모델을 활용한 태풍 예보 및 정보 제공 웹 서비스 “게셋거라(Stop Right There)”를 개발하였다.

2. 관련 연구

현대의 태풍 예측은 일반적으로 국내외 여러 장소에서 관측된 기상 데이터를 바탕으로 통계적·역학적 분석을 통해 수행한다. 대한민국 기상청은 통계적인 분석방법인 PC법, Climate법, Cliper법, 그리고 유사법에 더하여 국가기상슈퍼컴퓨터센터의 수치예보모델들을 활용한 역학적 분석을 진행하고, 최종적으로 기상 예보관들의 분석을 더하여 태풍 예측을 수행[2]하고 있다.

한편, 최근 딥러닝 기술의 발전에 힘입어, 국내외로 태풍 예측에 딥러닝 기술을 적용하여 예측

정확도를 향상하거나, 예측에 요구되는 자원을 최적화하는 연구가 활발히 진행되고 있다.

한국과학기술정보연구원은 시계열 위성영상 데이터를 사용한 딥러닝 기반 예측 모델 **PSIque**와 태풍의 대기 모델을 이용한 **LSTM** 기반 태풍 진로 예측 모델 **DeepTCI**^[3]를 개발하였다. **DeepTC**는 우리가 제안한 모델과 유사하게 기상 수치 예보 모델과 **best track** 데이터를 사용하였으며, 태풍의 미래 위도, 경도 값을 예측하여 6시간 후 태풍의 위치를 약 176km의 오차로 예측하였다.

인제대학교 연구진은 태풍의 이동경로, 중심 기압 및 이동속도에 대한 정보를 활용해 현재 진행 중인 태풍과 유사한 유사태풍을 분류하는 심층신경망 모델^[4]을 개발하였다. 해당 모델은 태풍의 미래를 예측하는 것이 아닌, 과거의 태풍 중 현재 태풍과 유사한 태풍을 검색하는 방식을 선택하였다. 이 연구에서 저자들은 데이터의 부족으로 인한 성능 하락을 근거로 관련 학습 데이터 구축의 필요성을 제시했다.

포항공과대학교 연구진은 위성영상 데이터로 생성적 적대 신경망(**Generative Adversarial Network, GAN**)을 학습시켜 미래의 태풍 위성영상을 생성하는 심층신경망 모델^[5]을 개발하였다. 해당 모델은 태풍의 12시간 후 경로를 68.7km 오차로 예측하였고, 18시간 후 경로 예측을 72.7% 정확도로 수행하였다. 그러나 생성 모델을 활용한 태풍 예측은 생성된 위성 영상에서 태풍의 눈을 찾는 방법으로, 모델이 태풍의 중심의 위치를 정확한 수치로 예측하는 것이 아니라 연구자가 결과물을 분석하여야 하는 만큼, 추가적인 인적 자원의 투입이 요구되며 정확성을 보장하기 어렵다.

이와 같이 기존 딥러닝 기반 태풍 예측 연구는 라벨로 사용된 **best track** 데이터의 구성에 따라 대부분 6시간이나 12시간과 같이 연구진이 사전에 정의한 시간이 지난 후의 태풍을 예측하도록

학습되었다. 이는 6시간 후의 태풍과 12시간 후의 태풍의 예측은 별도의 모델로 수행됨을 의미한다. **LSTM** 구조를 적용한 **DeepTC**는 6시간, 12시간, 18시간, 24시간 후와 같이 여러 시간대의 예측을 단일 모델로 수행할 수 있으나, 하나의 **LSTM** 셀이 6시간 후의 태풍을 예측하도록 학습되어 예측 시간이 6시간 단위로 나누어지며, 예측하고자 하는 시간이 미래일수록 순전파 과정에서 통과하는 **LSTM** 셀의 개수가 증가하여 연산량이 선형적으로 증가하는 한계가 있다. 한편 우리가 제안한 **TFDAN**은 예측하고자 하는 시간을 명시적으로 입력할 수 있으며, 예측 시간에 따른 연산량의 증가가 없어 더욱 효율적이고 유연하게 태풍의 미래를 예측할 수 있다.

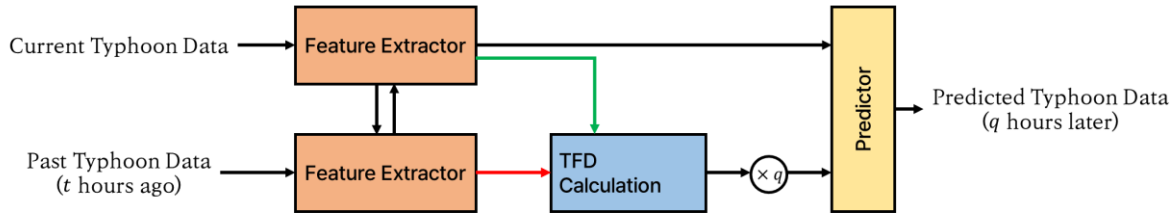
3. 데이터셋 구축

앞선 연구^[4]에서 지적인 것처럼, 높은 신뢰도를 갖는 심층신경망을 개발하기 위해서는 충분한 양의 고품질 데이터셋(**Dataset**)이 필수적이다. 데이터를 심층신경망 기법에 원활히 활용하기 위해서는 (1) 데이터의 양이 충분해야 하고, (2) 데이터에 누락된 값이 없어야 하며, (3) 데이터가 고정된 크기를 가져야 한다.

우리는 일본 기상청(**Japan Meteorological Agency, JMA**)에서 제공하는 **Best Track** 데이터와 유럽 중기 예보 센터(**ECMWF**)에서 제공하는 **ERA 5** 데이터를 기반으로 1977-2022년 사이에 발생한 태풍들의 데이터셋을 구축하여, 데이터 과학 분야에서 흔히 활용되는 **CSV** 형식으로 가공한 후 오픈소스 데이터 공유 웹사이트 **Kaggle**을 통해 공개¹² 하였다. 데이터셋은 **JMA best track** 데이터에서 취득한 태풍의 관측 시간(연, 월, 일, 시간), 위치(위도, 경도), 반경 등의 정보와, **ERA 5** 재분석 데이터에서 취득한 태풍 중심 및 반경 1,000km 8개 지점에 대한 온도, 습도, 기압, 구름의 양, 일조량, 풍속 및 풍향 등 10가지 기상 정보들로 구성된다. 데이터는 관측 시점당 6개의 **best-track** 정보와 태풍 중심 및 인근

¹ <https://www.kaggle.com/datasets/skyl7/asian-typhoon-dataset-1951-2022-rsmc-best-track>

² 현재 **best track** 데이터만 제공 중이며, **ERA 5** 기반 최종 데이터셋은 '23년 하반기 공개 예정



〈그림 2. 태풍 예측을 위한 시간적 특성 변화량 분석 신경망 구조〉

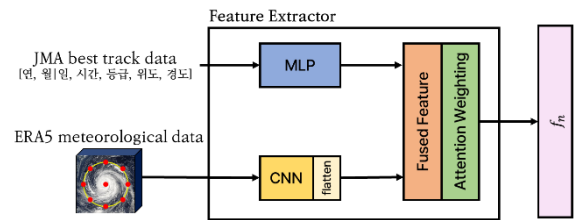
9개 지점에 대한 10가지 ERA 5 재분석 데이터를 더하여 최종적으로 96개의 값으로 구성되었다.

3.1. JMA best track Data

먼저, JMA best track 데이터에서 기상 관측 기준이나 태풍 분류 기준 등이 현재와 상이한 1977년 이전 데이터를 모두 제거하였다. 이후 남아있는 1977년부터 2022년까지의 1174개의 태풍 데이터 중, 2022년의 태풍 데이터 25개는 분리하여 모델의 성능을 평가하기 위한 테스트 데이터로 활용하였다. JMA best track 데이터에는 관측 시점에 따른 태풍의 중심 위치, 등급, 중심 기압, 최대 풍속, 태풍 중심으로부터 30노트와 50노트 각각의 풍속에 대한 반경이 포함된 관측 자료가 6시간 간격으로 기록되어 있는데, 태풍이 일본 기상청의 경계 영역으로 진입하면 관측 간격을 줄여 3시간 단위로 기록되어 있다.

3.2. ERA 5 재분석 데이터

태풍의 중심 데이터 외에 주위의 기상 상황까지 활용한 예측을 위해, 우리는 위에서 설명한 JMA best track에 더하여 태풍의 중심과 반경 1,000km의 8개 지점에서 ERA 5 재분석 데이터에서 태풍의 진행과 상관관계가 클 것으로 예상되는 온도, 습도, 기압, 구름의 양, 일조량, 10m와 100m 고도에서의 풍속과 풍향, 돌풍 정보를 취득하여 데이터셋을 추가하였다. 태풍으로부터 적절한 거리를 가진 지점에서 데이터를 추출하기 위해, JMA best track 데이터에 포함된 태풍의 영향 반경을 분석하였다. 그 결과, 태풍으로 인해 30노트 이상의 바람이 발생하는 영역은 평균적으로 태풍 주변 390km로부터 표준편차 188km



〈그림 3. 특성 추출 단계의 구조〉

이내의 구간임을 확인하였고, 태풍의 영향 반경을 최대 578km로 규정하였다. 이로부터 약 2배의 거리를 갖는 1,000km 지점에서 균등하게 태풍 주변 8개 방위의 ERA 5 데이터를 얻어 태풍 인근의 기상 상황을 고려한 예측을 수행할 수 있도록 하였다.

4. 방법론

우리는 과거와 현재 두 시점의 데이터와 6시간 단위로 주어지는 미래의 데이터로 학습되었음에도 미래의 태풍 데이터를 유연하게 예측하는 모델을 설계하기 위해 시간적 특성 변화량 분석 신경망(Temporal Feature Difference Analysis Network, TFDAN) 구조를 제안한다. 제안하는 구조는 〈그림 2〉와 같이 특성 추출 단계, 시간당 특성 변화량 계산 단계, 예측 수행 단계로 구성된다.

4.1. 특성 추출 단계

먼저, JMA best track 데이터로부터 얻은 정형 데이터 $x_b \in \mathbb{R}^5$ 와 ERA 5 데이터로부터 얻은 기후 데이터 $x_e \in \mathbb{R}^{3 \times 3 \times 10}$ 를 각각 다계층 퍼셉트론(Multi-layer Perceptron, MLP) 구조의 신경망과 합성곱 신경망(Convolutional Neural Network, CNN)에 투입하여 특성들을 추출한다. 구체적으로 x_b 는 2,048개의 노드를 갖는 은닉층 3개와

	6시간 후		12시간 후		18시간 후		24시간 후	
	RMSE	오차(km)	RMSE	오차(km)	RMSE	오차(km)	RMSE	오차(km)
best track	0.529	57.028	1.016	109.204	1.555	167.756	2.068	234.609
ERA 5	0.899	104.154	1.474	167.834	2.028	233.489	2.580	305.169
BT + ERA 5 (Circle 형태)	0.495	55.100	0.868	100.672	1.277	150.046	1.771	209.884
BT + ERA 5 (GRID 형태)	0.507	58.114	0.950	109.123	1.446	166.701	1.892	226.114

〈표 1. ERA 5 재분석 데이터 사용 여부와 형식에 따른 성능 비교〉

1,024개의 노드를 갖는 은닉층을 거쳐 1,024 차원의 특성으로 변환되며, x_e 는 입력 데이터의 높이와 너비는 유지하면서 채널의 개수를 각각 64개, 256개로 변환해주는 커널 크기 1의 합성곱 계층 두개를 거친 후, 커널 크기 3의 합성곱 계층을 축소되어 최종적으로 2차원의 특성으로 변환된다. 이렇게 얻어진 특성은 최종적으로 특성 벡터 $f \in \mathbb{R}^{1026}$ 로 융합된다.

그 다음, 신경망이 융합된 특성에서 중요한 요소들에 집중할 수 있도록 특성 벡터 f 에 어텐션 [6] 구조를 통해 구해지는 가중치 w 를 부여한다. 가중치 $w \in \mathbb{R}^{1026}$ 는 다음과 같이 정의된다.

$$w = \text{Sigmoid}(FC(f))$$

위 식에서 FC 는 완전 연결 계층을 의미한다.

특성 f 를 완전 연결 계층에 입력하고 시그모이드 함수를 이용하여 0과 1 사이의 값을 출력하게 한 후, 이를 다시 f 에 곱하여 중요한 특성에 더 높은 값을 부여하도록 학습시킨다.

4.2. 시간당 특성 변화량 계산 단계

앞서 정의한 특성 추출기를 통해 현재와 과거 두 시점의 태풍 정보로부터 특성을 추출한다. 연구 과정에서는 현재로 설정된 데이터로부터 6시간 전의 데이터를 활용하였다. 두 특성의 차를 구하고, 이를 두 정보를 측정된 시간의 차이로 나누어 시간당 특성 변화량을 계산한다.

$$\Delta f = \frac{f_{\text{present}} - f_{\text{past}}}{(h_{\text{present}} - h_{\text{past}})}$$

위 식에서 f 는 각각 현재와 과거 데이터에서

추출한 특성, h 는 각 데이터의 관측 시간을 의미한다.

4.3. 예측 수행 단계

미래 태풍의 예측은 앞서 얻은 현재 태풍의 데이터 f_{present} 와 태풍의 시간당 특성 변화량을 이용해 수행한다. 태풍의 시간당 특성 변화량에 예측하고자 하는 만큼의 시간 q 를 곱하고 이를 f_{present} 와 함께 예측 모델에 입력한다. 예측 신경망 P 는 현재 태풍의 데이터와 태풍의 변화량을 기반으로 태풍이 위도와 경도 방향으로 각각 몇 km 이동할 지 예측한다.

$$d_{\text{lat}}, d_{\text{lon}} = P(f_{\text{present}}, q\Delta f)$$

예측 결과 얻어진 태풍의 이동 거리를 위도와 경도로 변환하여 학습과 예측을 수행한다. 이는 위도와 경도는 지구의 곡률에 의해 값의 범위에 따라 의미하는 거리가 다르고, 적도 인근에서 위도가 불연속한 지점이 존재하기 때문이다.

4.4. 모델 학습 및 테스트

모델의 학습은 1977년부터 2021년까지 발생한 1149개의 태풍 데이터를 이용하였으며, 각 태풍의 한 시점의 위치를 각각 6시간 전의 데이터, 12시간 전의 데이터, 18시간 전의 데이터, 24시간 전의 데이터로부터 예측하도록 학습시켰다. 모든 경우에 예측 시점으로부터 6시간 이전의 데이터를 과거 데이터 삼아 예측하였다. 예를 들어, (6시간 전, 12시간 전)의 데이터로 현재 태풍의 상태를 예측하고, (12시간 전, 18시간 전)의 데이터로 현재 태풍의 데이터를 예측하는 방식이다.

	6시간 후		12시간 후		18시간 후		24시간 후	
	RMSE	오차(km)	RMSE	오차(km)	RMSE	오차(km)	RMSE	오차(km)
시간적 특성 변화량 분석 모델 (OURS)	0.529	57.028	1.016	109.204	1.555	167.756	2.068	234.609
심층 신경망 모델	0.613	69.013	1.273	143.67	2.091	235.453	2.596	311.794

〈표 2. 시간적 특성 변화량 분석 모델과 일반 모델의 성능 비교〉

학습에는 Adam 최적화 알고리즘을 사용하였고, 4096의 배치 크기를 이용하여 30 에포크 학습하였다. 매 학습 시 배치 내의 샘플을 섞어 모델이 과적합되는 것을 예방하였으며, 1⁻⁵의 L2 규제 값을 적용하였다. 실험 과정 전체에서 변수를 제거하기 위해, 모든 랜덤 시드값을 42로 고정하였다.

5. 비교 실험

태풍의 경로를 예측하는 최적의 심층신경망 모델과 데이터 구조를 찾아내고 우리가 제안한 모델의 성능을 입증하기 위해 다양한 형태로 모델과 데이터를 수정하며 비교 실험을 진행하였다.

5.1. ERA 5 재분석 데이터의 사용

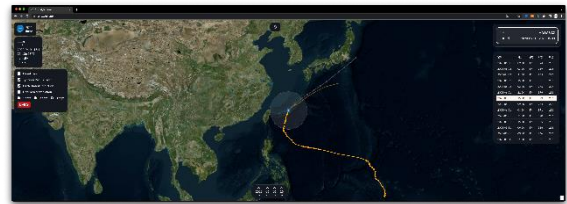
먼저 태풍의 기본적인 위치 정보만을 가지고 있는 best track 데이터만을 사용하여 태풍의 경로를 예측하는 경우와, ERA 5 기후 재분석 데이터를 함께 활용하는 경우의 성능을 비교하였다. 〈표 1〉에 나타난 실험 결과를 보면, ERA 5 데이터를 사용하지 않고 best track 데이터만 사용한 성능에 비하여 ERA 5 데이터를 원형으로 추출하여 추가한 경우 RMSE 값이 0.034, 약 6.4% 감소하였다. 추가로, ERA 5 데이터를 각 점이 750km의 간격을 갖는 5×5 크기의 격자 형태로 추출하여 ERA 5 데이터의 추출 형식에 따른 성능을 비교하였으나, 반경 1,000km에서 8개의 지점을 원형으로 추출한 경우가 가장 성능이 뛰어났다.

5.2. 시간적 특성 변화량 분석 모델 구조

우리가 제안한 시간적 특성 변화량 분석 모델

의 유효성을 입증하기 위해, 6시간 이전 데이터와 현재 데이터를 활용하여 6시간 후, 12시간 후, 18시간 후, 24시간 후의 태풍을 예측하는 일반적인 심층신경망 모델과 우리가 제안한 모델 간의 성능 비교 실험을 진행하였다. 〈표 2〉에 나타난 것처럼, 모든 경우에 우리가 제안한 모델이 더 좋은 성능을 보였으며, 특히 더욱 먼 미래의 예측을 수행할수록 우리가 제안한 구조의 성능이 더 우세한 것을 확인할 수 있다.

6. 태풍 예보 서비스 개발



〈그림 4. '게셋거라' 서비스 화면〉

우리는 연구 과정에서 개발한 모델을 활용한 태풍 예보 서비스 “게셋거라! (Stop Right There)”³를 개발하였다. 본 서비스는 실시간으로 전세계의 태풍 데이터를 수집하여, TFDAN 모델을 활용한 태풍의 향후 진로, 풍속, 등급 등 예측 정보를 제공한다. 사용자는 웹 페이지를 통해 태풍의 현재 정보와 예측 정보를 확인할 수 있으며, 원한다면 태풍 정보를 문자 메시지로 수신할 수 있다.

7. 요약 및 결론

우리는 매년 동아시아 지역에 막대한 피해를 입히는 태풍을 예측하기 위한 딥러닝 기반 초기 연구를 수행하였다. 이때, 기존 딥러닝 기반 태풍

³ <https://srts.web.app/>

예측 연구의 문제를 (1) 관련 연구를 위한 표준 데이터와 측정 기준의 부재, (2) 데이터에 의해 제약되는 예측 가능 범위로 정의하였다. 이러한 문제들을 해결하기 위해 먼저 csv 형태로 기계학습 연구에 적절하게 가공된 태풍 관련 데이터셋과 연구 결과를 평가하기 위해 딥러닝 기반 모델과 평가 코드를 구축하여 배포하였다. 또한, 두번째 문제를 해결하기 위해 불연속적인 데이터에서 학습되었음에도 가변적이고 연속적인 미래의 상태를 예측할 수 있는 시간적 특성 변화량 분석 신경망 구조와 이에 기반한 태풍 예측 모델을 제안하였다. 우리가 제안하는 모델은 6시간 후 태풍 위치 예측을 기준으로 기존 딥러닝 회귀 기법 기반 방법론인 DeepTC[3] 대비 약 69% 우세한 성능을 보였다.

8. 연구의 한계

본 연구는 연구진의 태풍에 관한 기반지식의 부족으로 인하여 태풍 인근 기후 정보 수집의 기준이 모호하고 미흡하였다는 한계가 존재한다. 또한, 태풍 인근 지점 이외의 외적 요소(다른 태풍, 전 지구적인 기후장 상황 등)를 고려하지 않아 24시간을 초과하는 태풍의 장기 예측 성능이 높지 않다. 향후 이러한 한계를 극복하여 기존의 역학적 분석 방법 대비 적은 자원으로 수행 가능한 높은 정확도의 딥러닝 연구가 진행되기를 바란다.

Acknowledgment

본 프로젝트를 진행하는 과정에서 따뜻한 관심과 핵심적인 조언을 아끼지 않고 제공해주신 최우석 교수님께 감사드립니다.

9. 참고 문헌

- [1] 기상청, “태풍이란,” [온라인].
<https://www.weather.go.kr/w/typhoon/basic/info1.do>.
- [2] 기상청, “태풍 예보법,” 기상청, [온라인].
<http://typ.kma.go.kr/TYPHOON/down/%C5%C2%C7%B3%B9%E9%BC%AD%C1%F5%BA%B8%C6%C71-7%C5%C2%C7%B3%BF%B9%BA%B8%B9%FD.pdf>.
- [3] Kim, Seongchan et al. “DeepTC: ConvLSTM Network for Trajectory Prediction of Tropical Cyclone using Spatiotemporal Atmospheric Simulation Data.” (2018).
- [4] 김연중, 김태우, 윤종성 그리고 김인호, “뉴럴 네트워크의 최적화에 따른 유사태풍 예측에 관한 연구,” %1 한국해양공학회지, 2019.
- [5] 유동현, “태풍의 경로 및 세기 예측을 위한 순환회선신경망 인공지능기술 개발,” 과학기술정보통신부, 2020.
- [6] A. V. e. al., “Attention Is All You Need,” %1 Conference on Neural Information Processing Systems, 2017.