최신 인공지능 기반 전지구 대기 대순환 모형들의 특성 및 장단점 비교 분석

김동훈[1], 문일주[1], 조민수[2]

[1]제주대학교 태풍연구센터
[2]한국과학기술정보연구원

인공지능 기술의 급격한 발달로 최근 2년여 만에도 전지구 대기 대순환 인공신경망모형들이 다수 개발되었다. NVIDIA의 FourCastNet (2022.02, 2023.06)은 AFNO와 SFNO를 사용한 Operator Learning 기술을 선보였으며, Huawei Cloud의 Pangu-Weather (2022.11)은 개념적으로 3차원 자료를 그대로 처리할 수 있는 3DEST와 시간 간격별 독립 학습을 통한 hierarchical temporal aggregation 기법을 소개하였다. Google DeepMind의 GraphCast (2022.12)는 차세대 신경망으로 주목받고 있는 Graph Neural Network(GNN)을 사용하면서 ERA5의 수평·연직 해상도와 동일한 해상도를 사용하였고, 상하이 AI 연구실에서 개발한 FengWu는 대기 변수들의 특성을 고려한 변수별 독립처리 기법을 개발함으로써 장기예측 성능을 높였다. Microsoft의 ClimaX (2023.07)는 동일 모형으로 장단기부터 기후 예측까지 모두 가능하며 전지구 및 지역뿐만아니라 모델의 고해상도화 까지 아우르는 통합 모델을 개발하였다. 또한, ECMWF의 AIFS (2023.10, 2024.01)는 GraphCast의 GNN에 Transformer의 특징을 융합한 attention-based GNN을 개발하였으며, 중국 Fudan 대학의 FuXi (2023.11)는 기존 10일 예측을 뛰어넘어 15일까지 예측하는데 0~5일과 5~10일, 10~15일로 나누어 독립 예측함으로써 10일 이상의 예측에서 현재 가장 좋은 성능을 보이고 있다. 더욱이 이 모든 모형들의 예측 성능이 ECMWF의 물리기반 수치모형인 IFS보다 앞서면서도 자원 사용량과 소요시간을 대폭 단축하는 기염을 토하고 있다.

본 연구에서는 추정 소스코드를 제공해 주고 있는 FourCastNet(v1 & v2)와 PanguWeather, GraphCast, FuXi, FengWu 등의 5가지 모델을 최근의 기상 자료를 사용하여 예측 성능을 비교해 보고, 모델 간의 특성과 장단점을 각각 비교 분석하였다. 이를 통해 각 모델의 예측 성능과 신경망 구조, 자료의 특성 등 다양한 측면에서의 차이점을 분석하고, 각 모델이 기상 예측 및 기후 모형 개발에 제공할 수 있는 기여와 한계를 가늠해 보고자 한다. 이러한 분석은 기상 및 기후 분야에서의 인공지능 기반 예측 모형의 발전 방향을 제시하고, 향후 연구 개발에 있어 중요한 시사점을 제공할 것으로 기대된다.

Key words: AI-based Global Weather Prediction Models, Machine Learning Weather Prediction(MLWP), Deep Learning in Meteorology, Data-driven Atmospheric Circulation Models

※ 이 연구는 한국과학기술정보연구원(KISTI)의 위탁연구개발과제와 2022년도 정부(과기정통부)의 재원으로 한국연구재단의 지원(No. RS-2022-00144325), 과학기술정보통신부 및 정보통신산업진흥원의 "2024년 고성능 컴퓨팅 지원"을 받아 수행하였음

특별세션 2 / HPCAI-AI-XR 기반 예측/분석 기술 2

기상 데이터 검색을 위한 AI 기반 자연어 및 음성 인식 시스템 개발

박준상[1,2], 김병연[1,2], 김태종[1,2], 조아름[1,2], 김소형[1,2], 이혜숙[1,2], 박채훈[3], 조진경[3], 주재걸[3], 서민준[3]

[1]국립기상과학원 인공지능기상연구과
[2]국립기상과학원 인공지능기상기술연구회
[3]한국과학기술원 김재철AI대학원

기상청의 실시간 모델 및 관측 자료는 방대한 양과 다양한 형태로 저장되어 기상 예보에 활용되고 있다. 방대한 자료에서 원하는 자료를 신속하게 찾아내지 못할 경우, 예보 생산 과정에서 상당한 시간이 소요된다. 따라서 신속하고 정확한 예보를 위해서는 필요한 기상 정보를 적시에 활용할 수 있는 효율적인 검색 시스템이 필수적이다.

본 연구는 방대한 기상 자료를 신속하고 정확하게 검색하여 예보 과정에서 소요되는 시간을 단축하는 기상-AI 검색기의 개발을 목표로 한다. 개발된 검색기는 기상청의 5세대 종합 기상 정보 시스템(COMIS-5)과 선진 예보 시스템의 자료를 학습하였고, 자연어 처리(NLP)와 음성 인식 기능을 포함한다. 자연어 처리는 mT5-base 기반 모델을 사용하여 기상 자료의 URL 및 극값을 효율적으로 검색한다. 음성 인식에는 Whisper-large 모델을 도입하여 기존 Whisper-medium 모델 대비 성능을 향상하였다. URL 검색을 위한 훈련 자료는 COMIS-5 웹페이지에서 키워드를 웹 크롤링하여 매핑한 데이터를 사용하였다. 검색 결과의 불균형 문제를 해결하기 위해 가짜 URL 데이터를 추가하였고, 기상 용어의 동의어 및 공백 처리의 학습자료를 추가하였다. 음성 인식 모델은 한국어 일반 음성 데이터와 예보 도메인 음성 데이터, TTS 합성 음성을 기반으로 훈련되었다.

자연어 처리 모델의 성능을 평가한 결과, COMIS-5에서의 MRR(Mean Reciprocal Rank)은 0.71, 선진 예보 시스템에서는 0.83으로 나타났다. 음성 인식 모델의 경우, 추론 시간은 약 0.39초로 기존 대비 약 38% 감소하였으며, 음절 단위 오류율은 10.49%에서 3.19%로 개선되었다.

Key words: 예보지원, 인공지능, 자연어 처리, 음성 인식, 검색기

※ 이 연구는 인공지능 기술지원 및 활용연구 중 AI 예보지원 및 활용기술 개발사업(KMA2021-00123)의 지원으로 수행되었습니다.

정지궤도 위성 영상을 활용한 인공지능 기반 기상 현상 분석 및 단기 예측 연구

김병연[1,2], 조아름[1,2], 이혜숙[1,2], 한창훈[3], 김수지[3], 박하명[3], 이건[4], 최민영[4], 신기정[4]

[1]국립기상과학원 인공지능기상연구과
[2]국립기상과학원 인공지능기상기술연구회
[3]국민대학교 소프트웨어융합대학
[4]KAIST 김재철AI대학원

기상 예보에서 중요한 요소 중 하나는 과거 데이터를 통해 유사한 기상 현상을 찾아 현재 기상과 비교하는 것이다. 그러나 기상 데이터는 방대하여 예보관들이 이를 효율적으로 활용하는 데 어려움을 겪고 있다. 본 연구에서는 정지궤도 위성에서 수집한 영상 데이터를 활용하여 유사한 기상 현상을 신속하게 검색하고, 인공지능을 적용해 향후 6시간 동안의 단기 예측을 제공하는 검색기를 제안한다. 본 연구에서는 다양한 기상 현상을 관측하는 정지궤도 위성인 COMS와 GK2A의 영상 데이터를 사용하였다. 제안된 검색기는 이미지 벡터화 모듈과 검색 및 예측 모듈로 구성되어 있다. 이미지 벡터화 모듈은 CNN과 RNN을 사용해 위성 이미지의 시공간적 특징을 추출하고 벡터화한다. 검색 및 예측 모듈에서는 NN-Descent 알고리즘을 통해 K-최근접 이웃(K-NN) 그래프를 최적화하여 유사 기상 현상을 빠르게 검색하고, 이를 기반으로 향후 6시간 동안의 기상 변화를 예측한다. 이 검색기는 사용자 인터페이스를 통해 구현되었으며, 위성 영상과 기상 지도를 포함한 다양한 기상 정보를 제공한다. 이를 통해 예보관들이 예보 작업을 보다 효율적으로 수행할 수 있도록 지원하며, 추가 기능을 통해 지속적으로 발전시킬 예정이다.

Key words: 예보지원, 유사사례, 위성 영상, 인공지능, 단기예측

※ 이 연구는 AI 예보지원 및 활용에 관한 연구 및 개발사업(KMA2021-00123)의 지원으로 수행되었습니다.

특별세션 2 / HPCAI-AI-XR 기반 예측/분석 기술 4

한국형모델(KIM)기반 AI-앙상블 예측시스템 소개

최원준, 신현철, 국현훈, 김은정, 신현철, 박종임, 하종철, 김동준

기상청 수치모델링센터 수치모델개발과

기상청은 기상예측 의사결정 지원 강화를 위해 2024년 10월부터 한국형모델(Korean Integrated Model, KIM) 기반의 AI(인공지능 모델, GraphCast) 앙상블 시스템을 준비중에 있다. 2021년 10월부터 운영되고 있는 한국형수치예보모델 기반 앙상블예측시스템의 각 멤버의 초기장을 25km로 만들어 입력으로 사용한 AI-앙상블은 총 26개의 앙상블 멤버(규준멤버 1개, 섭동 25개)를 활용하여 일 2회 (00 UTC, 12 UTC) 12일 예측을 수행한다.

인공지능 모델인 미국 구글 DeepMind 에서 개발된 GraphCast 는 Pangu-Weather, FourCastNet 보다 성능이 좋은 것으로 알려져 있으며, ECMWF 에서 실시간 표출되며 활용되고 있다. 따라서, 본 연구에서는 KIM 앙상블 초기장에 인공지능 모델(GraphCast)을 적용한 예측 결과에 대한 성능을 확인하고자 한다.

AI-앙상블 예측시스템의 성능은 2023년 여름철 충청지역 집중호우, 태풍 카눈 등 사례를 중심으로 검증지수를 비교 및 분석하였다. AI-앙상블의 예측 성능은 KIM 앙상블보다 전반적으로 우수하게 나타나 KIM 앙상블의 성능을 보완하는 효과를 보였다.

Key words: 인공지능, AI, 앙상블

※ 이 연구는 수치모델링센터 『수치예보 및 자료응용 기술개발』 과제 (KMA2018-00721)의 일환으로 수행되었습니다.

특별세션 2 / HPCAI-AI-XR 기반 예측/분석 기술 5

기상기후데이터와 뉴럴 네트워크 기반의 수위 데이터 이상치 탐지 – 오십천을 대상으로 –

추경수[1], 강우협[2], 김동희[2], 이혜준[2], 김현석[2], 김병식[3]

[1]강원대학교 AI기후재난기술융합연구소
[2]강원대학교 AI소프트웨어학과
[3]강원대학교 AI소프트웨어학과/방재전문대학원

최근 기후 변화로 인한 극한 호우로 인해 하천 범람 및 제방의 붕괴로 침수로 인한 피해가 증가하고 있다. 이에 따라 수위를 예측하여 침수를 예측하거나 효율적인 수자원 관리를 위해 기상기후데이터와 수위자료 등 수문 자료의 신뢰성이 확보되어야 하며 이를 위한 품질 관리는 매우 중요하다. 그러나 환경적, 물리적인 원인으로 이상치가 빈번하게 발생하고 있다. 따라서 수위 데이터를 목적에 맞게 활용하기 위해 이상치를 탐지하고 수위 데이터를 개선하는 것이 필요하다. 본 연구에서는 과거 침수 사상이 있는 삼척 오십천을 대상으로 딥러닝 뉴럴 네트워크 모델을 이용한 수위 추정 기술을 사용하고, 주변의 수위 및 강우 상황으로부터 목표 관측 지점의 현재 수위를 추정하고, 실제 관측 데이터와의 차이 정도로부터 이상도를 계산하였다. 이를 통해 이상 탐지의성능을 향상시킴으로써, 홍수 예측의 정확성을 높이고 방재 대응 체계를 강화하는 데 기여할 수 있을 것으로 기대된다.

Key words: 뉴럴 네트워크, 기상기후데이터, 이상치 탐지

※ 이 연구는 행정안전부 재난안전 공동연구 기술개발(2022-MOIS63-002(RS-2022-ND641012)) 사업의 지원을 받아 수행되었습니다.

특별세션 2 / HPCAI-AI-XR 기반 예측/분석 기술 6

동아시아 계절 강수 예측 개선을 위한 인공지능 기술 활용: Toward AI-aided APCC MME

이윤영[1], 정의현[2]

[1]APEC 기후센터 기후사업본부
[2]한국해양과학기술원 해양순환기후연구부

APEC기후센터(APCC)는 현재 11개국 15개 기후 예측 기관들의 독립 모델들을 최적 병합하는 다중모델앙상블(MME) 예측 시스템을 운영하고 있으며, 계절 시간규모에 대한 단정론적/확률론적 예보값을 가시화 혹은 디지털 데이터 형태로 제공하는 전지구 기후 정보 서비스를 수행 중에 있다. 2000년부터는 동아시아 지역 맞춤 다중모델앙상블 예측 정보를 생산해서 제공하는 등 세부 지역 기후 서비스로 활동 영역을 넓혀 왔다.

한반도를 포함하는 동아시아 지역은 대륙의 동안이라는 지리적 특성상 예측 성능이 현저히 떨어지며, 특히 물리 변수인 강수의 경우 다중모델앙상블 기술을 적용하더라도 그 시·공간적 분포에 대한 예측이 쉽지 않다. 기후변화가 가속화되는 상황에서 기후 관련 재난은 점점 가혹하고 빈번해지고 있으며, 최근 몇 년간 비정상적인 장마, 몬순 패턴으로 발생한 동아시아 여름철 홍수 및 가뭄 등이 그 실례라 할 수 있다. 장마 전선이나 몬순성 강우에 대한 계절 규모 예측 성능을 개선할 수 있다면 매년 반복되는 국가급 재난 상황에 대해 효과적으로 대응할 수 있을 것이다.

본 연구는 인공지능(AI) 기술을 활용해 다중모델앙상블 예측을 후보정함으로써 계절 강수 예측 성능을 개선할 수 있는지에 대한 가능성을 타진해 보았다. 이를 위해 딥러닝 기술을 이용해 1. MME에 참여하는 여러 앙상블 중 성능이 우수할 것으로 추정되는 개별 앙상블을 선별(SubSampling, SS)하고, 2. 관측 대비 역학 예측의 오차를 제거(Debiasing, DB)하는 두 가지 방식을 적용하였고, 전통적인 편차보정 방식(Quantile Mapping, QM)을 기준선으로 개선 여부 평가를 진행하였다. 그 결과, 2023년 12 계절 중에서 DB가 7개, SS가 4개 계절에서 현업 APCC-MME와 QM 성능을 넘어 최고 성능을 보여 동아시아 계절 강수 예측에 대한 딥러닝 기술 적용의 우수성/타당성 확인할 수 있었다.

본 발표에서는 인공지능을 활용한 계절 강수 예측 후처리 과정과 평가 결과를 공유하고, 각기 상이한 성능을 보이는 여러 인공지능 후처리 모형 결과를 병합하는 형태의 새로운 앙상블 접근법에 대해 논의하고자 한다.

Key words: 동아시아 강수, 계절예측 후처리, 인공지능