

*Bài báo khoa học*

# **Xây dựng công cụ ước lượng mưa từ độ phản hồi radar bằng công nghệ AI**

**Đỗ Hồng Hoạt<sup>1\*</sup>, Lê Ngọc Quyên<sup>1</sup>, Vũ Diệu Hồng<sup>1</sup>, Nguyễn Công Thành<sup>1</sup>**

<sup>1</sup> Đài khí tượng thủy văn khu vực Nam Bộ; dh.hoat@gmail.com; quyentccb@gmail.com; vudieuhong15@gmail.com; jackynguyen.kttv@gmail.com

\*Tác giả liên hệ: dh.hoat@gmail.com; Tel.: +84-908586859

Ban biên tập nhận bài: 12/1/2023; Ngày phản biện xong: 24/2/2023; Ngày đăng bài: 25/3/2023

**Tóm tắt:** Số liệu ra đa thời tiết là nguồn số liệu rất hữu hiệu được sử dụng để phân tích và dự báo thời tiết, đặc biệt là trong việc cảnh báo hạn cực ngắn phục vụ dự báo. Việc sử dụng hiệu quả và kịp thời số liệu radar sẽ mang lại ý nghĩa rất lớn. Do đó, nghiên cứu xây dựng công cụ ước lượng mưa bằng công nghệ trí tuệ nhân tạo (AI) trong việc sử dụng số liệu radar vào dự báo mưa hạn cực ngắn. Mô hình Convolutional Neural Networks (CNN) với cơ sở Rainet được dùng để tính toán và ước lượng mưa từ ảnh radar Nhà Bè, với dữ liệu các tháng mùa mưa từ tháng 5 đến tháng 11 năm 2019 và năm 2020 dùng để huấn luyện mô hình. Nghiên cứu đã áp dụng số liệu năm 2021 để kiểm nghiệm kết quả. Đánh giá các trường hợp cho thấy khả năng ước lượng mưa của công cụ là tương đối tốt về khả năng xảy ra mưa; về lượng mưa đa phần nhỏ hơn so với thực tế, tuy nhiên, sai số cũng không quá lớn.

**Từ khóa:** Ước lượng mưa; Convolutional Neural Networks (CNN).

## **1. Mở đầu**

Trong những năm gần đây, ngoài việc sử dụng mô hình dựa trên vật lý, các nhà dự báo và nhà nghiên cứu đã bắt đầu áp dụng các kỹ thuật AI rộng rãi hơn nhiều, khi chúng thể hiện sự hiệu quả trong nhiều ứng dụng, bao gồm hiệu chỉnh sai lệch sau mô hình, xử lý bộ dữ liệu lớn, giảm quá tải nhận thức và mở khóa thông tin chi tiết mới trong bộ dữ liệu lớn. AI và các công nghệ khoa học dữ liệu, đặc biệt là khoa học máy tính và khai thác dữ liệu, thu hẹp khoảng cách giữa dự báo mô hình số và thời gian thực bằng cách cải thiện độ chính xác. Các kỹ thuật AI cũng trích xuất thông tin không có sẵn từ các mô hình dự báo bằng cách kết hợp đầu ra của mô hình với các quan sát để hỗ trợ quyết định bổ sung cho người dự báo và người dùng [1–9].

Trên thế giới, có nhiều nghiên cứu trước đó sử dụng công nghệ AI vào dự báo, cảnh báo thời tiết nguy hiểm. [1] đã nghiên cứu máy học dự báo lượng mưa từ ảnh radar (Machine Learning for Precipitation Nowcasting from Radar Images). Các tác giả nghiên cứu một ứng dụng của Deep Learning (DL) cho vấn đề dự báo lượng mưa, độ phân giải cao (1 km × 1 km) trong thời gian ngắn (1 giờ) dự báo về lượng mưa. Các tác giả coi dự báo như một bài toán dịch từ ảnh sang ảnh và tận dụng sức mạnh của mạng nơ-ron tích tụ UNET phổ biến. Các tác giả thấy điều này hoạt động thuận lợi khi so sánh với ba mô hình thường được sử dụng: luồng quang học, độ bền và dự báo hiện tại HRRR một giờ của NOAA. Nghiên cứu chỉ ra rằng, dự báo độ phân giải cao là một công cụ thiết yếu cần thiết để thích ứng hiệu quả với biến đổi khí hậu, đặc biệt là đối với thời tiết khắc nghiệt. Nghiên cứu về sử dụng AI để cải thiện thời gian gian dự báo cho các hiện tượng thời tiết tác động mạnh [2]. Trong nghiên cứu, các tác giả chứng minh rằng việc áp dụng các kỹ thuật AI cùng với sự hiểu biết vật lý về môi trường

có thể cải thiện đáng kể kỹ năng dự báo đối với nhiều loại thời tiết có tác động mạnh. Họ cũng đưa ra các kỹ thuật AI có thể xử lý “dữ liệu lớn”, cung cấp thông tin chi tiết về các hiện tượng thời tiết có tác động lớn và nâng cao nhận biết về thời tiết có tác động lớn. [3] Nghiên cứu đã sử dụng AI để dự báo lượng mưa tháng và khí hậu tương lai cho lưu vực Bowen, một khu vực khai thác than quan trọng ở Queensland, Australia. Các dự báo thực hiện dựa trên các mô hình hoàn lưu chung (GCM) và chỉ ra rằng sẽ có sự thay đổi về thời gian và cường độ của lượng mưa ở lưu vực Bowen cùng với biến đổi khí hậu. Việc sử dụng mạng nơ-ron nhân tạo (ANN), ANN có thể được sử dụng để tạo dự báo hàng tháng trước 3 tháng. Những dự báo này có thể được cải thiện thông qua sự kết hợp tuyến tính có trọng số của các dự báo. [4] nghiên cứu động lực khí quyển, sử dụng AI trong dự báo thời tiết và khí hậu. Các tác giả nhận thấy có thể sử dụng mạng nơ-ron sâu để mô phỏng toàn bộ động lực của một tổng thể được đơn giản hóa mạnh mẽ mô hình hoàn lưu, cung cấp cả dự báo tốt về trạng thái mô hình vài ngày tới cũng như khí hậu ổn định trong dài hạn chuỗi thời gian.

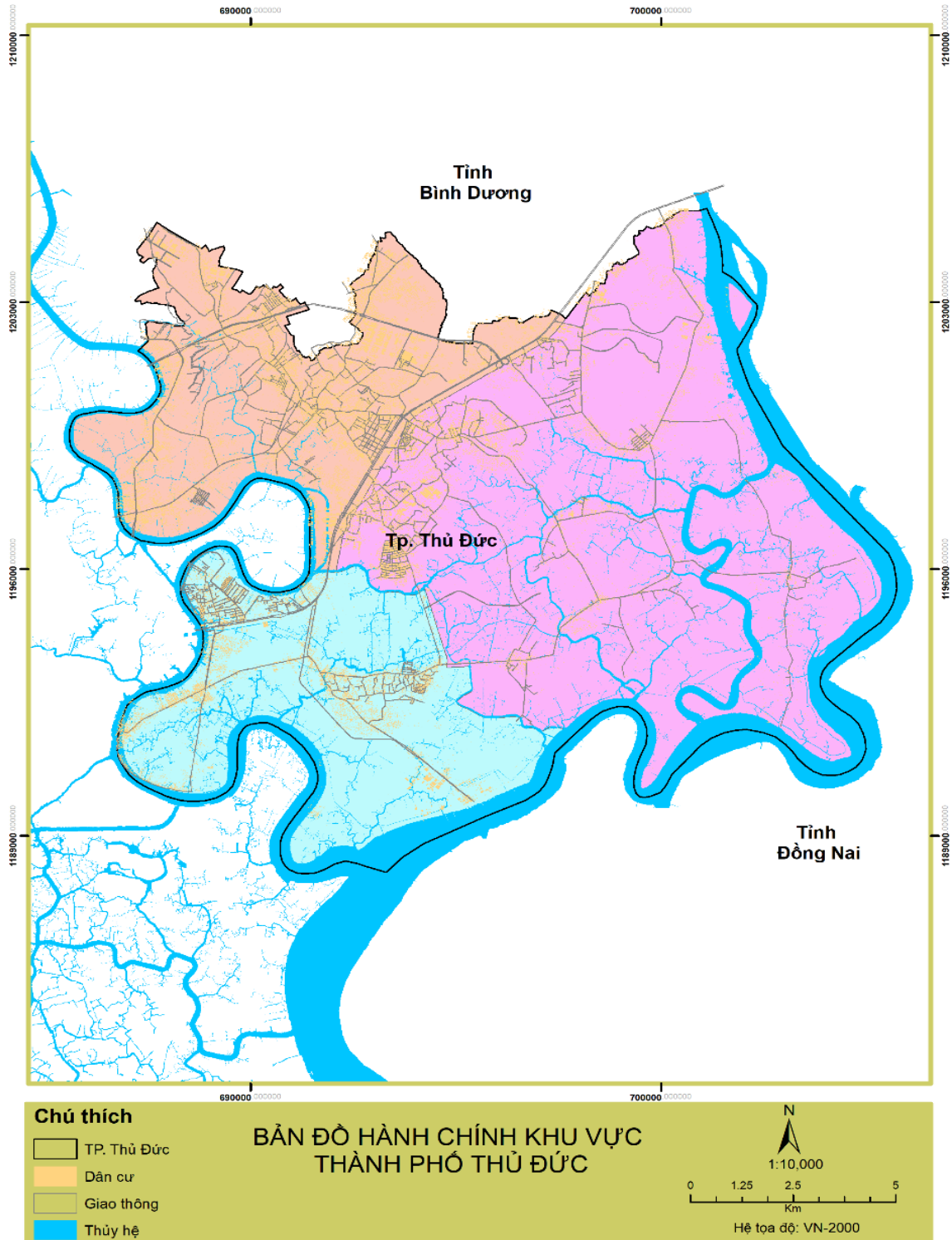
Tại Việt Nam, công nghệ AI cũng đang dần phổ biến trong nhiều lĩnh vực, và ngành Khí tượng Thủy văn cũng đã có những nghiên cứu ứng dụng công nghệ AI trong xử lý thông tin dữ liệu, quan trắc và dự báo, cảnh báo sớm. Đài Khí tượng cao không nghiên cứu, phát triển công cụ sử dụng mô hình học máy: *Logistic Regression Model (LRM)*, *Random Forest (RF)* và *Decision Tree Classifier (DTC)* để tăng cường độ chính xác của thông tin dự báo dông, sét đã được đưa vào dự báo nghiệp vụ từ năm 2022. Trung tâm Thông tin và Dữ liệu KTTV đã thực hiện nghiên cứu giải pháp ứng dụng AI để nhận dạng, hỗ trợ dự báo, cảnh báo một số hiện tượng khí tượng thủy văn nguy hiểm. Trung tâm Dự báo KTTV quốc gia có một số đề tài nghiên cứu xây dựng hệ thống AI ứng dụng trong dự báo xoáy thuận nhiệt đới ở Biển Đông và ảnh hưởng đến Việt Nam hạn đến 3 ngày; Nghiên cứu đổi mới công nghệ dự báo sóng biển, nước dâng do bão thời gian 24 giờ bằng kỹ thuật xử lý dữ liệu lớn và học máy và Nghiên cứu ứng dụng công nghệ số mới để dự báo định lượng mưa hạn cực ngắn cho khu vực trung du, miền núi Việt Nam. Đài KTTV khu vực Nam Bộ đã nghiên cứu xây dựng hệ thống giám sát, dự báo, cảnh báo ngập/ triều đô thị dựa trên nền tảng AI tại TP HCM; ứng dụng AI phục vụ dự báo mặn và thí điểm cho tỉnh Sóc Trăng [10].

Có nhiều phương pháp học máy để giải quyết bài toán xử lý hình ảnh phân loại hình ảnh (*Image classification (SVM, k-means, ...)*), phát hiện đối tượng (*Image detection (RCNN, Fast RCNN, Faster RCNN, YOLO, ...)*). Xây dựng công cụ ước lượng lượng mưa từ độ phản hồi Radar là một bài toán xử lý hình ảnh khá phức tạp đòi hỏi phải có sự tổng hợp kết quả của các phương pháp trên nên chúng tôi đã lựa chọn phương pháp CNN dựa trên Rainet có kiến trúc U-Net. Đây là một trong những phương pháp tiên tiến trong lĩnh vực xử lý ảnh có mô hình học sâu với nhiều lớp và có khả năng mở rộng. Nó giúp cho việc dự báo chính xác hơn và hiệu quả hơn. Chi tiết về phương pháp, quá trình thực nghiệm và kết quả sẽ được trình bày lần lượt ở các phần tiếp theo của nghiên cứu.

## 2. Số liệu và phương pháp nghiên cứu

### 2.1. Giới thiệu khu vực nghiên cứu

Trong nghiên cứu này, thành phố Thủ Đức, TPHCM được chọn làm trọng tâm nghiên cứu. Đối với những vấn đề cơ bản cần giải quyết để phát triển kinh tế xã hội cũng như khoa học kỹ thuật thuận lợi hơn đó là vấn đề giao thông và ngập úng, là hai vấn đề lớn. Tại TPHCM vấn đề ngập đang là vấn đề rất đáng quan tâm khi với mật độ dân cư cao, cũng như hệ thống hạ tầng dày đặc thì ngập lụt gây ảnh hưởng rất lớn tới đời sống của người dân, và khu vực thành phố Thủ Đức cũng là điểm ngập thường xuyên với những “rốn ngập” của thành phố. Vì vậy, việc lựa chọn thí điểm dự báo mưa lớn hạn cực ngắn nhằm phục vụ hệ thống cảnh báo sớm ngập lụt dựa trên nền tảng trí tuệ nhân tạo của đề tài cho thành phố Thủ Đức là một sự lựa chọn hợp lý và hữu ích; sau đó, khi hệ thống đã được kiểm nghiệm, ứng dụng cũng như đánh giá và cho kết quả tốt sẽ triển khai trên toàn bộ TP Hồ Chí Minh [11].



**Hình 1.** Bản đồ hành chính thành phố Thủ Đức.

## 2.2. Số liệu thu thập

Số liệu radar thời tiết Nhà Bè đặt tại huyện Nhà Bè, thành phố Hồ Chí Minh các tháng mùa mưa từ tháng 5 đến tháng 11 của giai đoạn 2019–2021, thời gian quét 10 phút/lần, được sử dụng để làm dữ liệu nghiên cứu. Số liệu các năm 2019, 2020 dùng để huấn luyện mô hình AI. Số liệu năm 2021 sử dụng để kiểm nghiệm kết quả. Độ phân giải mỗi tệp là 660×660.

Số liệu đánh giá: Số liệu mưa dự báo và quan trắc tại 3 điểm đo mưa tự động gồm trạm Cát Lái, trạm Thủ Đức, trạm Dĩ An vào các ngày 15/7/2021, 19/7/2021, 19–20/8/2021, 18–19/9/2021 và 21/9/2021.

### 2.3. Phương pháp nghiên cứu

Nghiên cứu sẽ sử dụng mô hình *Convolutional Neural Networks (CNN)* với cơ sở Rainet dùng để nghiên cứu và xây dựng công cụ ước lượng mưa từ số liệu radar. Mạng nơ-ron tích chập – CNN là một trong những mô hình Deep Learning tiên tiến giúp cho chúng ta xây dựng được những hệ thống thông minh với độ chính xác cao. CNN bao gồm tập hợp nhiều lớp khác nhau có thể học và trích xuất các đặc trưng quan trọng của hình ảnh một cách tự động và giúp cho việc dự đoán trở nên chính xác hơn. Thiết kế của nó được phát triển được ứng dụng rộng rãi trong cho các bài toán nhận dạng, phân tích hình ảnh và các bài toán dự báo mưa bằng một lượng lớn hình ảnh từ độ phản hồi radar [12].

#### 2.3.1. Phương pháp xây dựng mô hình AI

Mô hình AI được xây dựng theo kiến trúc mã hóa – giải mã (*encoder–decoder*) đối xứng nhau. Đặc trưng trong kiến trúc đối xứng giữa các lớp bên trái với lớp bên phải giúp ta phục hồi lại thông tin đã thất thoát tại các lớp max pooling vì nếu như ta sử dụng lớp lấy mẫu tăng lên (*upsampling layer*) ngay từ lớp cuối cùng ở phần mã hóa thì thông tin của dữ liệu ban đầu bị mất đi rất nhiều.

- Trong đó phần mã hóa (*encoder*) dùng để giảm kích thước đầu vào của lớp dữ liệu qua đó trích lọc đặc trưng để tìm ra đặc trưng của dữ liệu bằng việc sử dụng các lớp tích chập (*convolutional layer*) và các lớp tổng hợp tối đa (*max pooling layer*).

- Phần giải mã (*decoder*) dùng để phục hồi lại kích thước ban đầu của lưới dữ liệu sử dụng các lớp lấy mẫu tăng lên (*upsampling layer*) và lớp tích chập (*convolutional layer*).

Mô hình có 20 lớp tích chập (*convolutional layer*), 4 lớp tổng hợp tối đa (*max pooling layer*), 4 lấy mẫu tăng lên (*upsampling layer*) và 2 lớp bỏ qua (*dropout layer*) và 4 lớp kết nối bỏ qua (*skip connections*).

#### 2.3.2 Phương pháp đánh giá

Nghiên cứu lựa chọn các chỉ số cho đánh giá gồm: ME, MAE, RMSE, Bias, Far, PC [16–19].

Đánh giá sai số trung bình: F–Dự báo, O–Quan trắc

- Sai số trung bình ME

$$ME = \frac{1}{n} \sum_{i=1}^n (F_i - O_i) = \bar{F} - \bar{O} \quad (1)$$

- Sai số trung bình tuyệt đối

$$MAE = \frac{1}{n} \sum_{i=1}^n |F_i - O_i| \quad (2)$$

- Sai số quân phương

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (F_i - O_i)^2} \quad (3)$$

Đánh giá đối với dự báo pha: Hits (H) = dự báo có và quan trắc có; Misses (M) = dự báo không và quan trắc có; False alarms (F) = dự báo có và quan trắc không; Correct negatives (CN) = dự báo không và quan trắc không.

- Điểm sai số (bias score): Bias score = (H+FA)/(H+M)

Bias score là thước đo tần số tương đối giữa mưa dự báo và quan trắc.

- Tỷ phần dự báo phát hiện sai: FAR = F/(H+ F)

Giá trị tối ưu FAR (False Alarms Ratio) = 0

- Độ chính xác: PC (Percentage Correct) = (H+ CN)/ (M+ F+ H+CN)

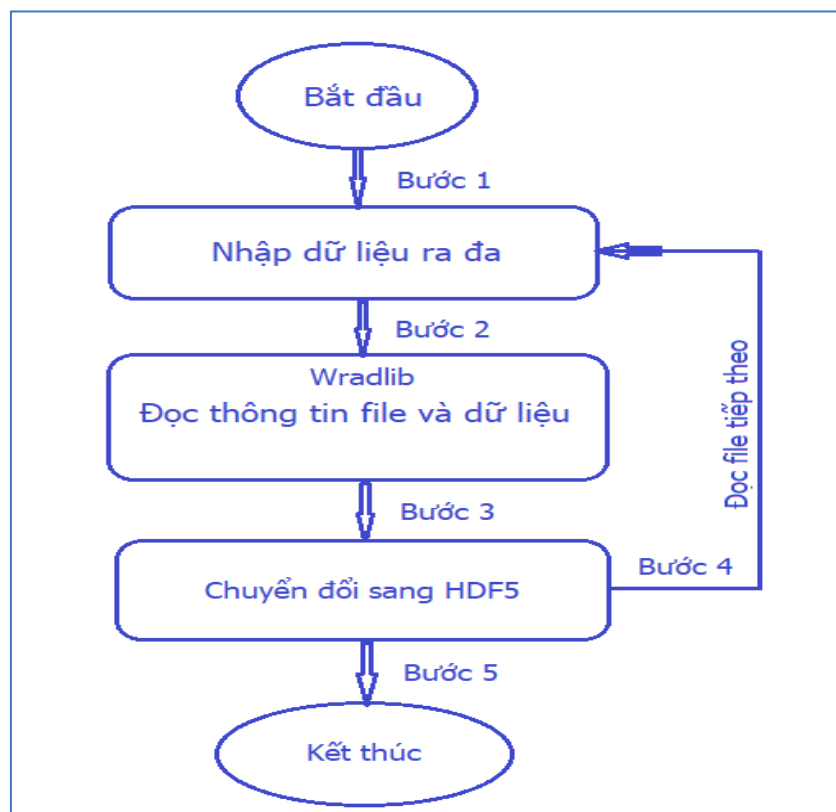
Giá trị tối ưu FAR (False Alarms Ratio) = 1

#### 2.4. Quá trình thực nghiệm mô hình

##### 2.4.1. Chuyển đổi và sắp xếp dữ liệu ra đa RAW ở định dạng HDF5

Sau khi thu thập số liệu ra đa Nhà Bè, sẽ thực hiện chuyển đổi dữ liệu ra đa về định dạng HDF5, tiến hành cài đặt và sử dụng thư viện mã nguồn mở Python Wradlib, đây là thư viện phát triển và sử dụng riêng cho việc phân tích dữ liệu ra đa, được rất nhiều trung tâm trên thế giới sử dụng bởi tính hiệu quả và ổn định. Sau đó, tiến hành viết một chương trình python để đọc các file RAW và chuyển đổi về định dạng HDF5. Quy trình thực hiện chuyển đổi theo các bước sau:

- Bước 1: Chương trình đọc file dữ liệu radar tại thư mục được chỉ định.
- Bước 2: Dùng thư viện Wradlib đọc file radar.
- Bước 3: Chuyển đổi dữ liệu sang định dạng HDF5.
- Bước 4: Chuyển về bước 1 với file ra đa tiếp theo, thực hiện lại các bước 2, 3.
- Bước 5: Sau khi chuyển đổi tất cả các file ra đa sang HDF5 sẽ kết thúc quá trình.



**Hình 2.** Quy trình chuyển đổi dữ liệu ra đa về định dạng HDF5.

##### 2.4.2. Thu thập và lập dữ liệu mưa quan trắc tự động

Do mỗi loại trạm đo mưa có những cấu trúc dữ liệu khác nhau nên chúng tôi đã xây dựng chương trình chuyển đổi đối với từng loại dữ liệu. Ngôn ngữ Python được sử dụng để xây dựng chương trình. Đối với các trạm thuộc dự án “Đo mưa nhân dân VRAIN”, phần mềm sẽ lấy số liệu theo phương thức HTTP qua website của dự án mỗi 10 phút. Theo dõi các thư mục chứa file data được truyền về mỗi máy chủ riêng biệt của từng dự án. Mỗi khi có file dữ liệu được truyền về chương trình sẽ đọc file dữ liệu thô rồi thông qua các bước chuyển đổi bên dưới để đưa về dữ liệu chuẩn.

Khi thu thập đầy đủ các số liệu từ các trạm tự động và tiến hành kiểm soát số liệu thì chúng tôi sẽ tiến hành thiết kế và xây dựng cơ sở dữ liệu My SQL.



2.4.3. Các bước chạy mô hình

Huấn luyện mô hình AI với các bộ dữ liệu radar trong quá khứ để thu được bộ trọng số. Bộ trọng số này sẽ được dùng để làm trọng số cho mô hình AI.

Mô hình AI lấy bốn lưới dữ liệu radar liên tiếp làm các kênh đầu vào riêng biệt (t-30, t-20, và t-10 phút và t – trong đó t là thời gian của hiện tại) để dự báo thời điểm t+10 phút. Mỗi lưới chứa 660×660 giá trị; đối với mỗi ô, giá trị đầu vào là lượng mưa lấy từ sản phẩm phản hồi vô tuyến của radar. Có 5 mức độ phân giải gần như đối xứng cho cả bộ giải mã và bộ mã hóa, sử dụng các dữ liệu lượng mưa ở độ phân giải đầu vào là (x, y), ở một nửa độ phân giải (x/2, y/2), tại (x/4, y/4), tại (x/8, y/8) và tại (x/16, y/16). Cuối cùng, lớp đầu ra của độ phân giải (x, y) với hàm kích hoạt cung cấp lượng mưa dự đoán trong mỗi ô lưới trong t+10 phút. Sau khi dự báo ước lượng lượng mưa tại thời điểm t+10, chúng tôi sử dụng các giá trị t-30, t-20, t-10 và giá trị t làm đầu vào để dự báo ước lượng lượng mưa cho thời điểm t+20 quá trình được lặp lại để dự báo ước lượng lượng mưa cho thời điểm tiếp theo [13–14].

2.4.4. Quy trình tối ưu hóa

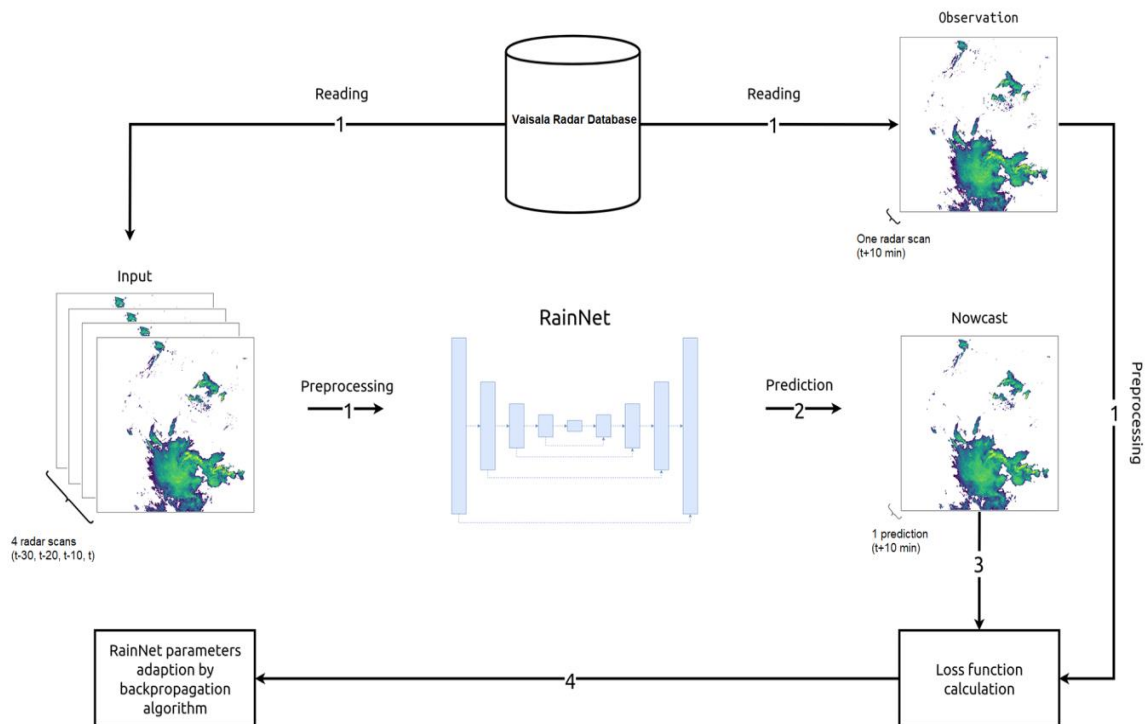
Nghiên cứu đã tối ưu hóa các tham số này bằng cách sử dụng một quy trình mà chúng tôi hiển thị một lần lặp trong hình bên dưới.

Đầu tiên, chúng tôi đọc một mẫu dữ liệu đầu vào bao gồm các lần quét radar tại thời điểm t-30, t-20, t-10 phút và t (trong đó t là thời điểm quan trắc) và một mẫu lượng mưa dự báo được tại thời điểm t+10. Bước thứ hai, mô hình AI đưa ra dự báo ước lượng lượng mưa dựa trên dữ liệu đầu vào. Bước thứ ba, tính toán một hàm mất mát (loss function) thể hiện độ lệch giữa dự báo và quan trắc. Chúng tôi sử dụng hàm logcosh như sau:

$$Loss = \frac{\sum_{i=1}^n \ln(\cosh(now_i - obs_i))}{n}; \cosh(x) = \frac{1}{2}(e^x + e^{-x}) \quad (4)$$

now<sub>i</sub> và obs<sub>i</sub> được phát sóng và quan sát tại vị trí thứ i, tương ứng; cosh là hàm cosin hyperbolic; n là số ô trong lần quét radar.

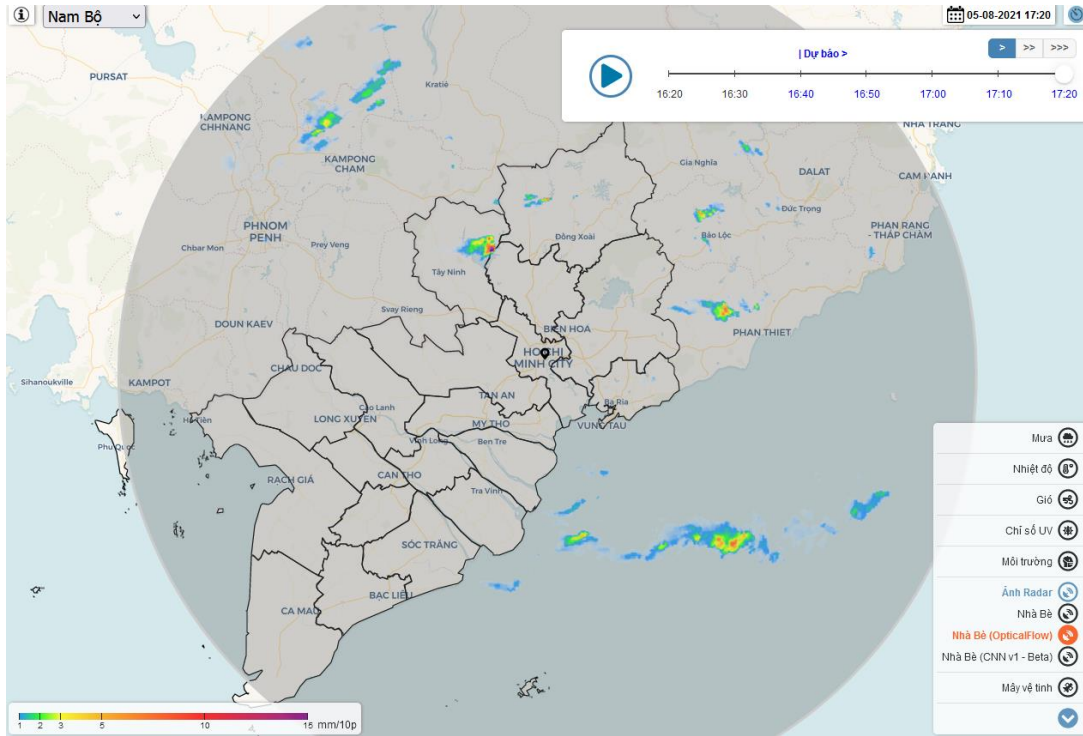
Bước bốn, cập nhật các tham số mô hình bằng hàm lan truyền ngược Adam optimizer. Minh họa một bước lặp lại của quy trình tối ưu hóa được thể hiện trong Hình 3 [13–14].



Hình 3. Minh họa một bước lặp lại của quy trình tối ưu hóa.

### 3. Kết quả và thảo luận

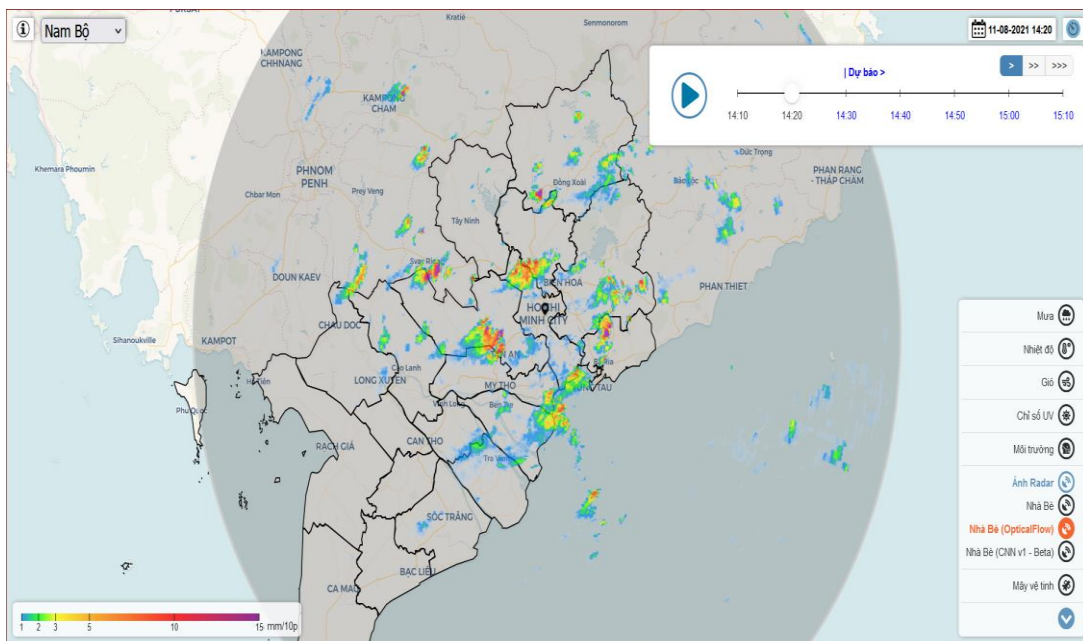
Sau khi nghiên cứu và xây dựng công cụ, chúng tôi đã hoàn thiện và thử nghiệm thành công cũng như đưa vào sử dụng công cụ ước lượng mưa từ số liệu radar bằng công nghệ AI. Công cụ hiển thị ảnh Radar Nhà Bè (Hình 4) với bên trên là thanh thời gian gồm khoảng thời gian thực tế và các khoảng thời gian dự báo để người sử dụng có thể lựa chọn các khoảng thời gian. Bên dưới là thang đo cường độ mưa đơn vị là mm/10p, cường độ mưa ước lượng được chia chi tiết theo thang màu để người dùng dễ dàng sử dụng.



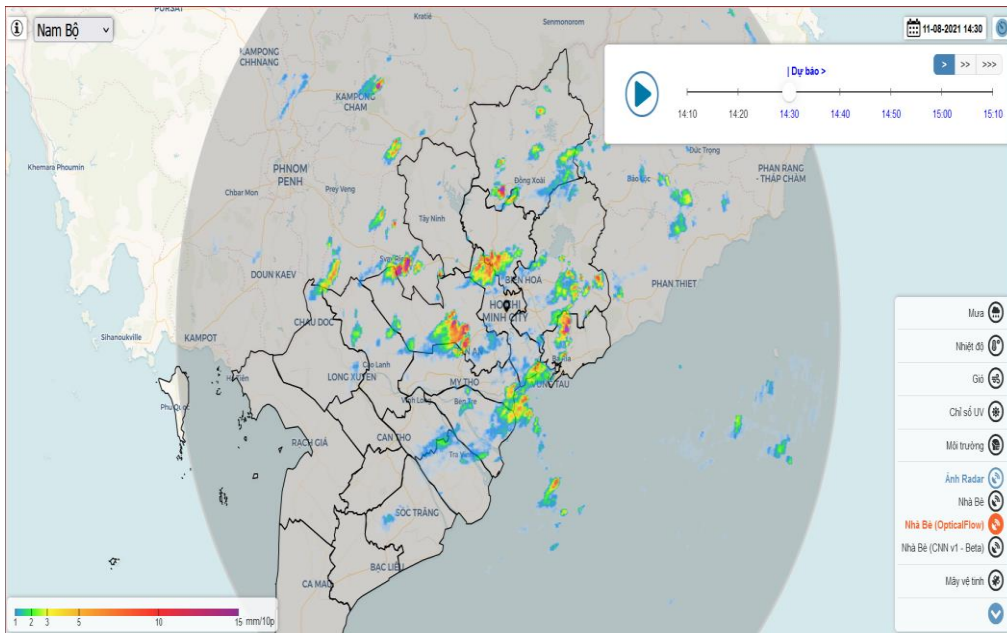
Hình 4. Giao diện công cụ ước lượng mưa từ ảnh radar bằng công nghệ AI.

#### 3.1. Kết quả thử nghiệm

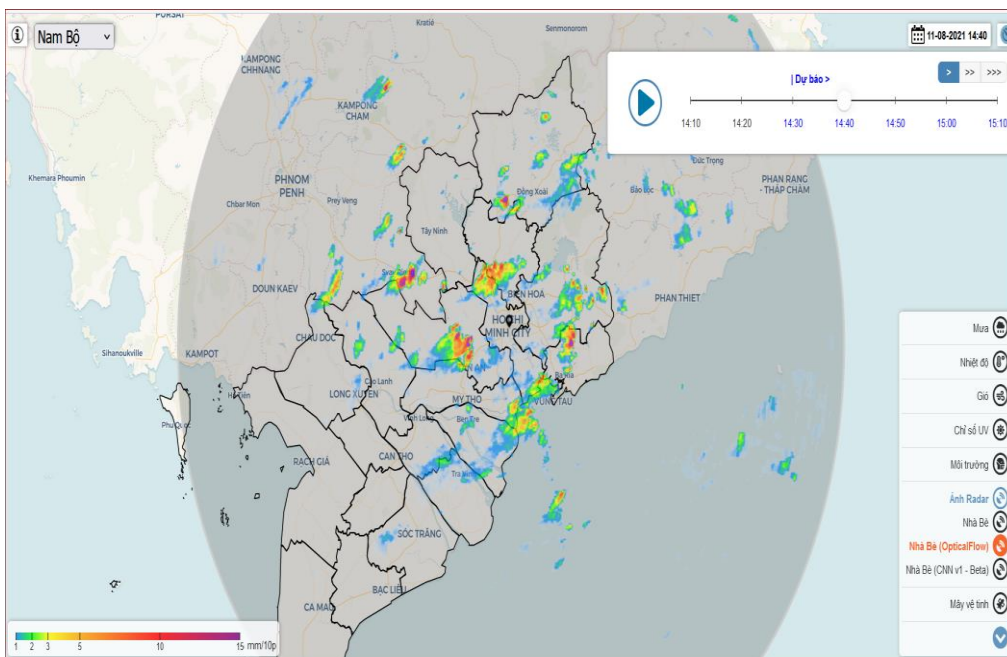
Trường hợp ngày 11/8/2021



Hình 5. Ảnh radar lúc 14h20 ngày 11/8/2021.



**Hình 6.** Ảnh radar dự báo lúc 14h30 ngày 11/8/2021.



**Hình 7.** Ảnh radar dự báo lúc 14h40 ngày 11/8/2021.

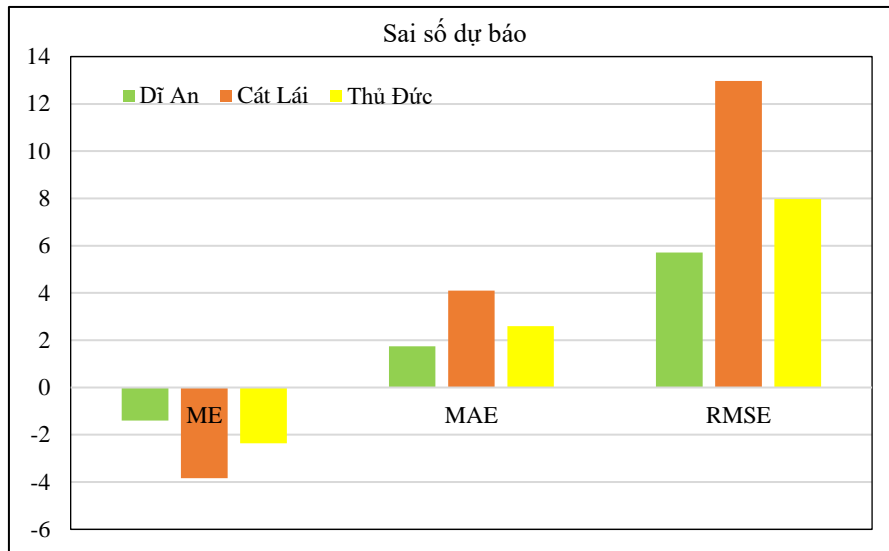
Trường hợp ngày 11/8/2021 (Hình 4–Hình 6) cho thấy rõ công cụ đã đưa ra được ước lượng mưa và thông qua các hình ảnh dự báo có thể thấy rằng vùng phản hồi di chuyển khá chậm khi vị trí các vùng mây cũng như lượng mưa ước lượng là ít thay đổi. Lượng mưa ước lượng cũng khá sát khi ước lượng ở Củ Chi–TPHCM khoảng 5 mm/10 phút và vùng mây đối lưu tồn tại khoảng 40 phút.

### 3.2. Kết quả đánh giá

Sau khi xây dựng thành công công cụ ước lượng mưa và đưa vào sử dụng, nghiên cứu sử dụng kết quả tại 3 điểm đo mưa tự động gồm Dĩ An, trạm Cát Lái, trạm Thủ Đức và các ngày 15/7/2021, 19/7/2021, 19–20/8/2021, 18–19/9/2021 và 21/9/2021. Tiến hành đánh giá bằng việc tính toán các chỉ số ME, MAE, RMSE, Bias, Far, PC, kết quả được phân tích trong Hình 8 và Hình 9.

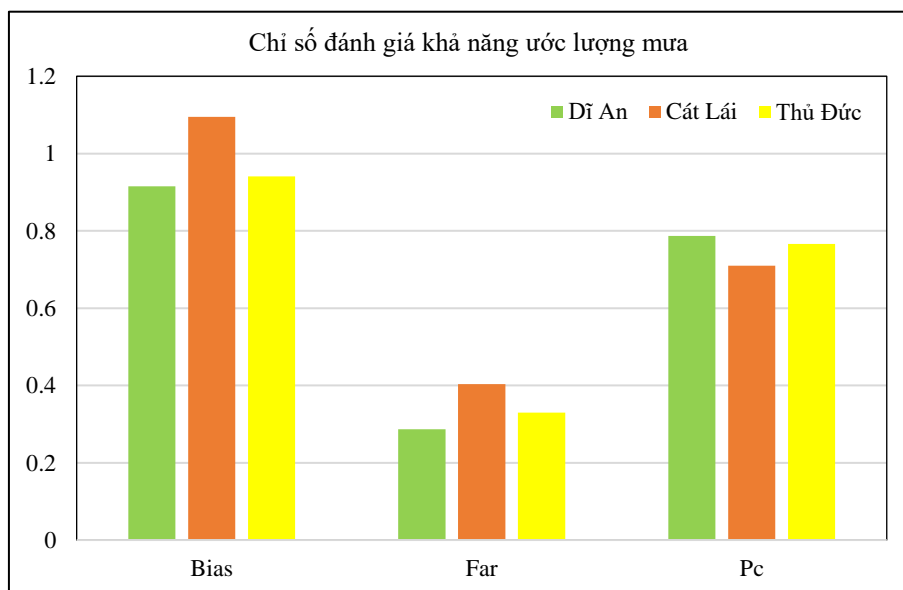


Hình 8 cho thấy các giá trị ME < 0, cho thấy rằng ước lượng mưa đều nhỏ hơn so với quan trắc thực tế, nhưng cũng nhỏ hơn không nhiều. Giá trị MAE dao động từ 1,75–4,09, sai số trung bình cũng không quá lớn và cũng không chênh lệch nhiều giữa các trạm, lớn nhất tại trạm Thủ Đức và nhỏ nhất là trạm Dĩ An. Giá trị RMSE của cả 3 trạm dao động từ 5,71–12,96, và chênh lệch giữa MAE và RMSE cũng không nhỏ (4–8,8) nên có thể thấy rằng ước lượng mưa ở cả 3 trạm đều có những trường hợp có sai số lớn.



**Hình 8.** Các chỉ số đánh giá sai số trung bình.

Theo dõi Hình 9 thấy rằng giá trị Bias giữa các trạm không chênh lệch nhau nhiều và đều dao động quanh giá trị 1 (từ 0,92–1,07), (Bias tối ưu = 1), do vậy mà xác suất ước lượng khá tốt, trạm Dĩ An và Thủ Đức có Bias < 1 nên ước lượng mưa số lần ít hơn so với thực tế, tuy chênh lệch là khá nhỏ; còn trạm Dĩ An có Bias > 1 nên số lần ước lượng mưa nhiều hơn thực tế nhưng cũng chênh lệch nhỏ. Giá trị Far từ 0,28–0,38 cho biết khả năng ước lượng mưa cũng không có nhiều dự báo không, khoảng chưa tới 28%–38% trên tổng số lần dự báo. Chỉ số Pc cho biết độ chính xác trong ước lượng có mưa hay không có mưa, và Pc lần lượt của các trạm Dĩ An 0,79, Cát Lái 0,72, Thủ Đức là 0,77. Các giá trị Pc cũng tương đối tốt và cho thấy khả năng ước lượng chính xác mưa hay không mưa tại 3 trạm Dĩ An, Cát Lái, Thủ Đức lần lượt là 79%, 72% và 77%.



**Hình 9.** Các chỉ số đánh giá dự báo pha.

#### 4. Kết luận

Nghiên cứu đã xây dựng được công cụ ước lượng mưa từ ảnh radar thời tiết bằng công nghệ AI. Công cụ cho thấy được sự hữu ích khi đã dự báo được lượng mưa bằng phương pháp sử dụng công nghệ AI và cụ thể là mô hình RainNet để tính toán và ước lượng mưa từ ảnh radar Nhà Bè với dữ liệu dùng các tháng mùa mưa từ tháng 5 đến tháng 11 với khoảng thời gian 10 phút. Số liệu 2019, 2020 dùng để huấn luyện mô hình và số liệu năm 2021 sử dụng để kiểm nghiệm kết quả; độ phân giải mỗi tệp là 660×660. Có thể thấy rằng, việc ước lượng lượng mưa từ ảnh radar thời tiết là rất ý nghĩa đối với dự báo cực ngắn và giúp dự báo viên rất nhiều trong việc đưa ra các cảnh báo kịp thời và chính xác hơn. Bên cạnh việc ước lượng được lượng mưa dự báo thì người sử dụng còn có thể theo dõi được sự dịch chuyển của những khối mây thông qua lượng mưa ước lượng và cảnh báo mưa lớn. Do vậy, công cụ khá hữu ích cho việc cảnh báo mưa lớn để giúp cung cấp số liệu và thông tin cho phục vụ chống ngập. Qua đánh giá thấy rằng khả năng ước lượng mưa của công cụ là tương đối tốt, tuy nhiên thường ước lượng nhỏ hơn so với thực tế và độ lớn sai số sẽ tăng dần theo bước thời gian ước lượng mưa.

Tuy nhiên công cụ vẫn có hạn chế nhất định đó là mới dự báo được lượng mưa trong thời gian ngắn. Mô hình huấn luyện AI đòi hỏi một lượng lớn dữ liệu để học và thời gian để huấn luyện khá lâu. Nghiên cứu sẽ tiếp tục phát triển và mở rộng để hoàn thiện hơn.

**Đóng góp tác giả:** Xây dựng ý tưởng nghiên cứu: Đ.H.H., L.N.Q.; Thu thập số liệu và xây dựng công cụ: Đ.H.H., N.C.T.; Đánh giá kết quả: Đ.H.H., V.D.H.; Viết bản thảo bài báo: Đ.H.H., V.D.H.; Chỉnh sửa bài báo: Đ.H.H.

**Lời cảm ơn:** Các tác giả xin gửi lời cảm ơn tới Đài Khí tượng Thủy văn khu vực Nam Bộ đã tạo điều kiện để thực hiện bài báo này. Nghiên cứu cũng là một phần kết quả thuộc đề tài “Nghiên cứu, xây dựng và triển khai thử nghiệm hệ thống cảnh báo sớm ngập lụt đô thị dựa trên nền tảng trí tuệ nhân tạo tại thành phố Hồ Chí Minh”.

**Lời cam đoan:** Tập thể tác giả cam đoan bài báo này là công trình nghiên cứu của tập thể tác giả, chưa được công bố ở đâu, không được sao chép từ những nghiên cứu trước đây; không có sự tranh chấp lợi ích trong nhóm tác giả.

#### Tài liệu tham khảo

1. Agrawal, S.; Barrington, L.; Bromberg, C.; Burge, J.; Gazen, C.; Hickey, J. Machine learning for precipitation nowcasting from radar images. arXiv preprint arXiv:1912.12132, 2019.
2. McGovern, A.; Elmore, K.L.; Gagne, D.J.; Haupt, S.E.; Karstens, C.D.; Lagerquist, R.;... Williams, J.K. Using artificial intelligence to improve real-time decision-making for high-impact weather. *Bull. Am. Meteorol. Soc.* **2017**, *98(10)*, 2073–2090.
3. Abbot, J.; Marohasy, J. Using artificial intelligence to forecast monthly rainfall under present and future climates for the Bowen Basin, Queensland, Australia. *Int. J. Sustainable Dev. Plann.* **2015**, *10(1)*, 66–75.
4. Scher, S. Artificial intelligence in weather and climate prediction. On Friday 12 June 2020 at 10.00 in Vivi Täckholmsalen, Svante Arrhenius väg, 20, 2020.
5. Marzban, C.; Stumpf, G. A neural network for tornado prediction based on Doppler radar-derived attributes. *J. Appl. Meteor.* **1996**, *35*, 617–626. Doi:10.1175/1520-0450(1996)035 <0617: ANNFTP>2.0.CO;2.
6. Mellit, A. Artificial intelligence technique for modeling and forecasting of solar radiation data: A review. *Int. J. Artif. Intell. Soft Comput.* **2008**, *1*, 52–76. doi:10.1504/IJAISC.2008.021264.

7. McCandless, T.C.; Haupt, S.E.; Young, G.S. A regime-dependent artificial neural network technique for short-range solar irradiance forecasting. *Renewable Energy* **2016**, 89, 351–359. doi:10.1016/j.renene.2015.12.030.
8. Wang, F.; Mi, Z.; Su, S.; Zhao, H. Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters. *Energies*, **2012**, 5, 1355–1370. doi:10.3390/en5051355.
9. Baldwin, M.; Kain, J.; Lakshmiarahan, S. Development of an automated classification procedure for rainfall systems. *Mon. Wea. Rev.* **2005**, 133, 844–862. doi:10.1175/MWR2892.1.
10. <http://vnmha.gov.vn/chi-dao-dieu-hanh-103/tong-cuc-khi-tuong-thuy-van-day-manh-ung-dung-tri-tue-nhan-tao-trong-linh-vuc-khi-tuong-thuy-van-14112.html>.
11. Quyên, L.N. và cs. Nghiên cứu, xây dựng và triển khai thử nghiệm hệ thống cảnh báo sớm ngập lụt đô thị dựa trên nền tảng trí tuệ nhân tạo tại Thành phố Hồ Chí Minh. Báo cáo tổng kết đề tài NCKH cấp Thành phố, 2022.
12. Wu, J. Introduction to convolutional neural networks. National Key Lab for Novel Software Technology. Nanjing University, China, 2017, 5(23), 495.
13. Ayzel, G.; Scheffer, T.; Heistermann, M. RainNet v1. 0: a convolutional neural network for radar-based precipitation nowcasting. *Geosci. Model Dev.* **2020**, 13(6), 2631–2644.
14. <https://github.com/hydrogo/rainnet>
15. <https://gmd.copernicus.org/articles/13/2631/2020/>
16. Tân, P.V. Phương pháp thống kê trong khí hậu, NXB Đại học quốc gia Hà Nội, 2005.
17. Toàn, N.T và cs. Đánh giá khả năng dự báo mưa lớn của mô hình WRF do hình thể không khí lạnh kết hợp với gió đông trên cao cho khu vực Trung Trung Bộ. *Tap chí Khoa học ĐHQGHN: Các Khoa học Trái đất và Môi trường* **2018**, 34(1S), 132–136.
18. Nga, N.T.; Thanh, C.; Hung, M.K.; Tiến, D.Đ. Đánh giá kĩ năng dự báo mưa định lượng từ mô hình quy mô toàn cầu và khu vực phân giải cao cho khu vực Bắc Bộ. *Tap chí Khí tượng Thủy văn* **2021**, 730, 79–92.
19. Đánh giá chất lượng dự báo mưa định lượng của mô hình WRF cho khu vực Việt Nam. *Tap chí Khí tượng Thủy văn* **2022**, 738, 1–11.

## **Building Rainfall Estimation Tool From Radar Reflectivity Using Artificial Intelligence Technique**

**Do Hong Hoat<sup>1\*</sup>, Le Ngoc Quyen<sup>1</sup>, Vu Dieu Hong<sup>1</sup>, Nguyen Cong Thanh<sup>1</sup>**

<sup>1</sup> Southern Regional Hydrometeorological Center; dh.hoat@gmail.com; quyentccb@gmail.com; vudieuhong15@gmail.com; jackynguyen.kttv@gmail.com

**Abstract:** Radar data is very effective in nowcasting and warning of heavy rain and floods for major cities and river basins. Effective and appropriate use of radar data is very important. Therefore, this study builds a rain estimator using Artificial Intelligence (AI) Technology in using radar data to forecast extremely short-term rainfall. Convolutional Neural Networks (CNN) model with Rainet basis is used to calculate and estimate rain from Nha Be radar images, with data of rainy season months from May to November 2019 and 2020 used to train the model. The study applied the 2021 data to test the results. Evaluation of the cases shows that the tool's ability to estimate precipitation is relatively good in terms of the probability of rain occurrence; rainfall is mostly smaller than reality, however, the bias is not too large.

**Keywords:** Rainfall Estimation; Convolutional Neural Networks (CNN).