

# 增强热带气旋路径预测的改进变换器网络

1<sup>st</sup> Nguyen Van Thanh

*Faculty of Electronics and Telecommunications*  
*University of Engineering and Technology, VNU*  
Hanoi, Vietnam  
felicenguyen2001@gmail.com

2<sup>nd</sup> Nguyen Dang Huynh

*Faculty of Information Technology*  
*University of Engineering and Technology, VNU*  
Hanoi, Vietnam  
nguyendanghuynh1804@gmail.com

3<sup>rd</sup> Nguyen Ngoc Tan

*Faculty of Electronics and Telecommunications*  
*University of Engineering and Technology, VNU*  
Hanoi, Vietnam  
tan.nguyen@vnu.edu.vn

4<sup>th</sup> Nguyen Thai Minh

*Faculty of Electronics and Telecommunications*  
*University of Engineering and Technology, VNU*  
Hanoi, Vietnam  
minhnt@hnm.edu.vn

5<sup>th</sup> Nguyen Nam Hoang

*Faculty of Electronics and Telecommunications*  
*University of Engineering and Technology, VNU*  
Hanoi, Vietnam  
hoangnn@vnu.edu.vn

**摘要**—风暴是一种极端天气。因此，预测风暴的路径对于保护人类生命和财产极其重要。然而，风暴预报非常具有挑战性，因为风暴轨迹经常发生变化。在这项研究中，我们提出了一种使用 Transformer 网络改进的深度学习方法来预测未来 6 小时内风暴的移动轨迹。用于训练模型的风暴数据来自美国国家海洋和大气管理局 (NOAA) [1]。仿真结果表明，所提出的方法比传统方法更准确。此外，所提出的方法更快且更具成本效益。

**Index Terms**—路径预测，热带风暴，深度学习，变压器模型

## I. 介绍

热带气旋被认为是极端天气现象。大型风暴（风速达到 63 公里/小时或更高）可能会对全球沿海地区的人们生命和财产造成重大损失 [2]。每年，大约有 40 到 50 个热带低压区（在赤道附近形成）发展成为风暴。准确及时的风暴路径预测使当局能够采取适当的预防措施，最大限度地减少风暴造成的损害。然而，由于风暴运动的复杂性和全球气候变化的影响，预测风暴轨迹面临着许多困难。

风暴路径预测技术在近年来取得了显著进展。然而，大多数当前的预测方法本质上是统计性的 [3]。具体来说，美国国家飓风中心 (NOAA) 目前使用了诸如以下几种预测方法：

- 全球预报系统 (GFS) [4] 和飓风天气研究与预报模型 (HWRF)；
- 地球静止轨道 operational environmental satellite (GOES) 系统 [5]；
- 陆地和海上天气传感器网络。

首先，GFS 和 HWRF 模型利用基于观测数据和当前天气状况的算法和方程计算并预测风暴运动。然而，GFS 的准确性依赖于输入数据的质量和精度，如风力强度、气压和气象信息等。这在预测快速发展的不可预见因素（如风暴的发展和转变）时带来了挑战。

第二个系统，GOES，收集关于风暴轨迹和强度的连续数据。通过实时监测风暴变化，NOAA 可以提供更新的信息和预报。然而，这取决于卫星的数据采集和传输能力，这些能力可能会受到恶劣天气或技术问题的影

响。对偏远海洋地区的有限卫星覆盖也带来了挑战。

最后，美国国家海洋和大气管理局使用一个气象传感器网络来收集诸如压力、风速、温度和湿度等因素的数据。这种方法需要在大面积范围内部署和维护复杂的传感器和数据采集设备系统，在偏远或孤立的海洋区域这可能会很困难。

上述方法受到大气系统复杂性和非线性以及求解复杂方程中计算误差的限制。

最近，一些基于深度学习的风暴预测方法出现了，例如循环神经网络 (RNN)、长短期记忆 (LSTM) 网络 [6] 和门控循环单元 (GRU)，它们能够有效地存储时间信息。之前，卷积神经网络 (CNN) 已被应用，因为它们有能力检测并学习数据中的空间特征，这使得它们适合根据风暴路径的空间特性预测风暴轨迹 [7]。然而，这些方法仍然存在一些限制，例如难以处理长期的风暴轨迹。模型通常只考虑每个时间步长的空间特征，而没有捕捉到沿轨迹的数据点之间的空间相互作用。当风暴轨迹规模较大或需要同时处理多个特性和数据源时，这尤其成为一个问题。传统的深度学习模型如 RNN、GRU 或 CNN 可能难以有效地扩展以应对如此复杂任务和大规模数据集 [9]。

Transformer 模型能够学习诸如湿度、气压、位置等气象变量之间的非线性关系，并且具有处理不同长度序列的能力，这使其在提高风暴轨迹预测准确性方面具有巨大潜力。因此，在本文中，我们提出使用 Transformer 模型来解决上述方法的局限性。

本文的结构组织如下：

原始 Transformer 模型和我们提出的版本的架构及操作流程分别在第 II 节和第三节中给出。最后，在第四节中，我们将使用所提出的 Transformer 模型预测风暴轨迹的结果与其他深度学习模型（如 LSTM）以及美国国家气象服务目前使用的预报方法进行了比较。

## II. 变压器模型

Transformer 模型是由 Vaswani 等人于 2017 年提出的用于处理序列数据的神经网络架构 [10]。它利用了多头注意力机制和前馈神经网络，并在各种自然语言处理任务中表现出高效性。因此，在本研究中，我们应用 Transformer 模型进行热带气旋轨迹预测，以发挥其优势——特别是通过注意力层学习湿度、压力、温度、风速等因素之间非线性关系的能力。

此功能使模型能够捕捉不同天气变量之间的相关性并提高预测准确性。此外，Transformer 模型在选择合适的模型大小以匹配输入数据的复杂性和计算需求方面具有灵活性。如图 5 所示，Transformer 架构主要由两部分组成：**编码器**和**解码器**。编码器和解码器均由多层组成，每一层都包含一个注意层和一个前馈层。编码器用于从输入数据中提取特征，而解码器则利用这些特征生成预测 [11]。

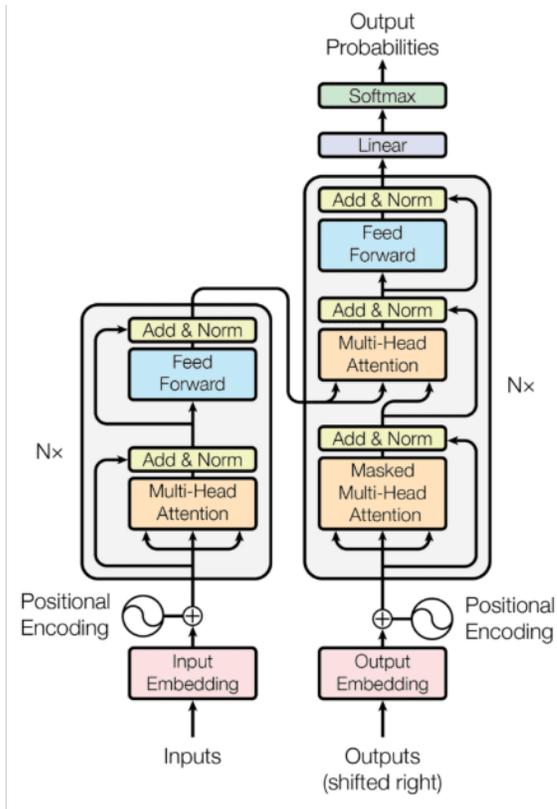


图 1. 原始 Transformer 架构概述，包括编码器和解码器组件。

几项最近的研究还将改进的 Transformer 模型应用于诸如干旱预测 [12] 以及追踪和预报热带气旋强度 [13] 等任务。在 [12] 中，作者提出了一种具有四个编码器和解码器层的深度 Transformer 模型。结果显示，Transformer 在长期预测性能上优于传统的 LSTM 模型。发表于 2023 年 7 月的研究 [13] 提出了一种用于同时追踪风暴轨迹并预报风暴强度的 Transformer 网络。在此工作中，作者使用了中国的台风数据进行实验，结果优于基于 RNN 的模型。

### III. 提议的结合坐标网格的 TRANSFORMER 模型

基于上述对 Transformer 架构的分析，在本节中，我们提出了一种结合了 Transformer 模型和坐标网格以提高准确性并减少方法复杂性的模型，如下所示。

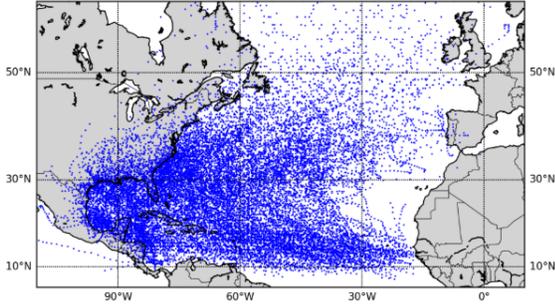


图 2. 1944 年至 2022 年的风暴发生点

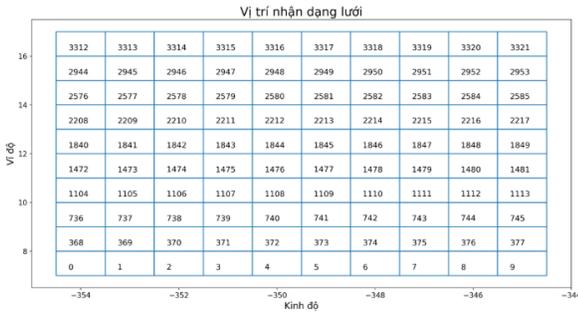


图 3. 一个分辨率约为 1° 经度 × 1° 纬度的坐标网格。

#### A. 网格模型

图??显示了用于训练和测试模型 [14] 的 Unisys Weather Dataset 记录的热带气旋的经纬度点。该图说明了在给定大气条件下风暴轨迹的典型移动。在一个精细网格上放置了经纬度点，以减少截断误差，同时允许模型捕捉大尺度模式并更精确地表示小尺度现象。这种设置非常适合使 Transformer 能够最优地捕获风暴轨迹的复杂性。

在本文中，我们提出了将 Transformer 模型与坐标网格模型相结合的训练方法，如 [15] 所示。具体来说，Transformer 学习风暴轨迹从一个网格单元移动到另一个网格单元的过程。图??所示的网格模型总共包含 23,533 个可识别的网格点。

#### B. 数据集

我们使用美国国家飓风中心 (NOAA) 提供的数据库 [1]。该数据集包括 1852 年至 2022 年大西洋盆地热带气

旋的信息。每个数据点包含每 6 小时的风暴中心位置 (纬度和经度)、风速 (以节为单位) 以及最低中心压力。在移除有故障或缺失数据的年份后，最终的数据集包括 1944 年至 2022 年的飓风，总计 982 次飓风，包含 22,545 条数据记录 (每 6 小时一条)。

从风暴的经度和纬度数据中，我们计算出诸如距离和移动方向等额外特征以支持轨迹预测。当前位置与未来 6 小时预报位置 (以英里为单位) 之间的距离是使用地理 OPY 库计算得出的，该库允许我们在给定坐标的情况下测量两点间的大地线距离。

类似地，风暴移动的方向是通过两个连续位置之间的角度  $\beta$  计算得出的，如下所示 [16]:

$$\angle\beta = \arctan\left(\frac{\sin(\Delta\lambda)\cos(\phi_2)}{\cos(\phi_1)\sin(\phi_2) - \sin(\phi_1)\cos(\phi_2)\cos(\Delta\lambda)}\right) \quad (1)$$

其中  $\angle\beta \in [0, 360]$ 。如果  $\beta$  是负数，则加上 360 以确保其为正数。 $(\lambda_1, \phi_1)$  和  $(\lambda_2, \phi_2)$  分别表示当前和下一个 6 小时步长的风暴经度和纬度，以及  $\Delta\lambda = \lambda_2 - \lambda_1$ 。

我们还计算与风暴在坐标网格上的位置对应的网格标识符 (网格 ID) 为:

$$\text{gridID} = \lfloor \lambda - \lambda_{\min} \rfloor \Phi + \lfloor \phi - \phi_{\min} \rfloor \quad (2)$$

其中  $\lfloor x \rfloor$  表示地板函数， $\lceil x \rceil$  表示最接近整数函数，而  $\Phi = \phi_{\max} - \phi_{\min}$ 。

#### C. 数据处理

Transformer 模型需要顺序输入数据。然而，风暴轨迹的长度各不相同。为了解决这个问题，我们在每个风暴序列的末尾应用零填充以标准化它们的长度。我们数据集中最长的风暴序列包含 96 个时间步，因此我们将所有序列填充到固定长度 100。

因此，训练数据集包含 98,200 个样本，每个样本有 5 个特征：风速、压力、距离、移动方向和网格标识符。

接下来，数据被分割成每个包含 13 个元素的段。前 12 个元素用作输入，第 13 个<sup>th</sup> 元素的网格标识符用作预测标签。最后，数据集按照 85:15 的比例分为训练集和测试集以评估模型的性能。

#### D. 提出的变压器模型架构

图??说明了本文提出的 Transformer 模型架构。数据转换为序列形式后，每个序列包含 12 个时间步长，

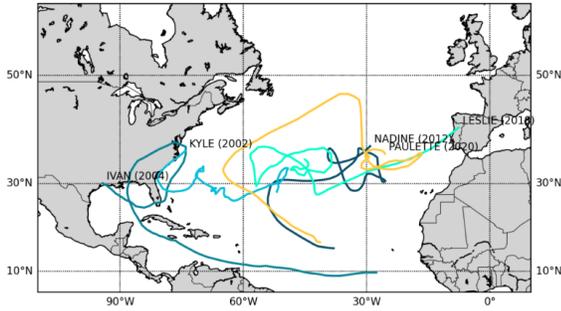


图 4. 持续时间最长的 5 个风暴的轨迹

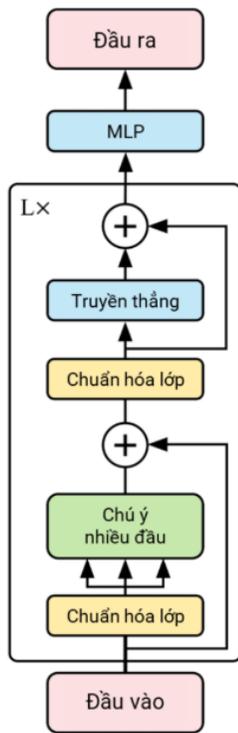


图 5. 提议的变压器模型架构

每个步骤表示为一个 5 维向量。在一个序列内，注意力层使每个位置能够观察其他位置，从而通过关注序列中更重要的部分重新编码当前输入。特别地，这种注意力机制通过多头注意力得到增强，使得每个位置可以同时关注多个相关特征。

这些变换后的元素然后通过具有高斯误差线性单元 (GELU) 激活函数 [17] 的前馈层，这提升了向量在后续层中的表示。

数据通过三个编码器层进行处理以获得更好的特征表示。最后，输出被送入一个具有两个子层的多层感知机 (MLP) 输出层：其中一个包含 12 个神经元，最

终的一个包含单个神经元。使用的激活函数为隐藏层的 ReLU 和输出层的 tanh。最后一个神经元代表预测的网格标识符。

在最终激活函数中使用双曲正切函数 (tanh) 而不是 ReLU 或 Sigmoid，使模型能够输出值范围为  $[-1, 1]$ ，从而更有效地预测风暴的移动方向 [15]。

图??和图??分别展示了在各个时期上的训练和测试损失及准确率。

#### IV. 结果与评估

提出的 Transformer 模型在一个配备 Tesla T4 GPU 的 Google Colab 环境中进行了训练。训练持续了大约 6 分钟，共 100 个 epochs 以最小化损失。该模型使用均方误差 (MSE) 损失函数、Adam 优化器进行训练，并通过准确率作为指标进行评估。

训练后，模型在测试集上实现了 0.0086 的均方误差，准确率为 0.78，如表 I 所示。

图??和??分别展示了所提出的 Transformer 模型对飓风伊万 (2004 年) 和飓风德尔塔 (2020 年) 的风暴路径预测的成功案例。

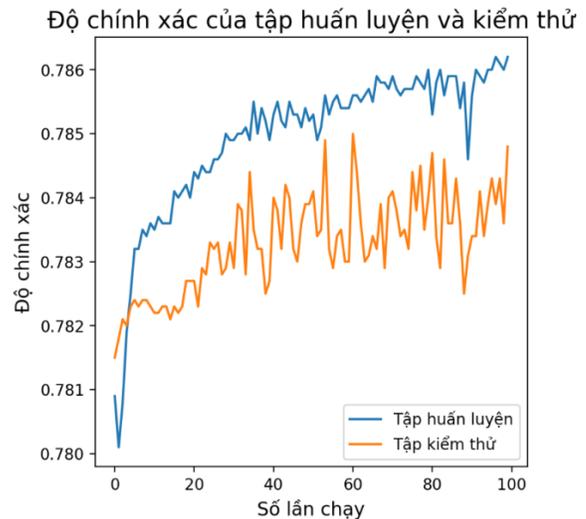


图 6. 图表显示每次运行后训练集和测试集的准确性

模型以 6 小时为间隔预测风暴移动。如图所示，预测的网格位置与实际风暴路径非常吻合，展示了模型的准确性。

国家飓风中心 (NHC) 的绩效指标作为 1993 年《政府绩效与成果法案》(GPRA) 的一部分被报告为 [14]。当前，NHC 的指标追踪大西洋盆地内所有热带气旋

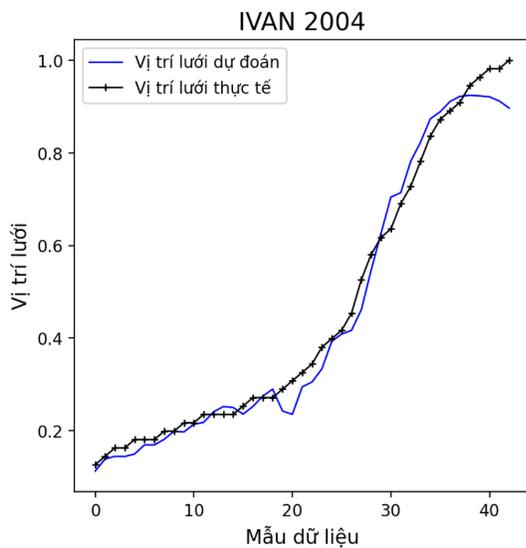


图 7. 飓风伊万 (2004 年) 的预测轨迹

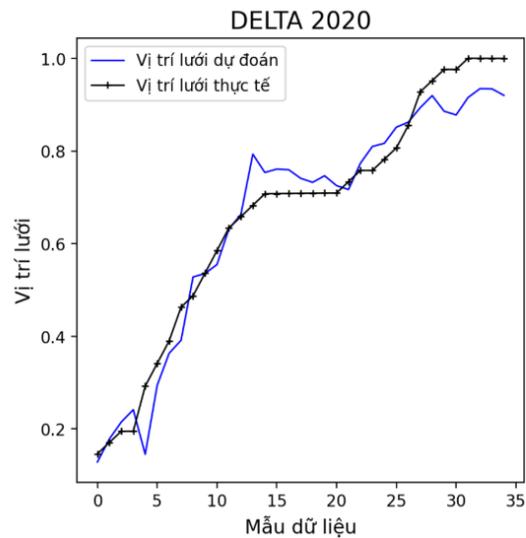


图 8. 飓风 DELTA (2020 年) 的预测轨迹

48 小时内风暴位置和强度的平均年度预测误差。尽管 NHC 提供从 12 小时到 120 小时间隔的预报，但 48 小时预报对于应急管理者和准备工作特别重要。由于风暴特征的自然变化，年度误差可能每年显著波动。

表 I

LSTM 模型与提出的 TRANSFORMER 模型的准确性比较

模型	均方误差	准确性
LSTM Model [6]	0.0160	0.685
Proposed Transformer Model	0.0086	0.783

图 9 显示，所提出的基于 Transformer 的预测方法和 NOAA 的方法均超过了 GPRA 设定的最低准确度要求。此外，使用坐标网格的提出的 Transformer 模型比现有的 NHC 方法具有更高的准确性。

此外，NHC 目前使用的统计动力模型通常需要几个小时才能生成一个预测结果，这使用了世界上一些最先进的超级计算机。相比之下，像所提出的 Transformer 这样的深度学习模型可以显著更快地提供准确的预测。

然而，所提出的 Transformer 模型在风暴轨迹预测中仍面临若干挑战。例如，如果调节不当或数据集不平衡，该模型可能会出现过拟合。这会导致轨迹预测失败。此外，Transformer 架构计算密集且复杂，尤其是在大规模数据集上训练时。这为将模型部署到资源受限的设备上带来了困难。



图 9. NOAA、GPRA 标准与提出的 Transformer 模型之间的预测准确性比较。

## V. 结论

在这项研究中，我们实现了一种使用 Transformer 模型结合基于网格的坐标映射系统预测连续 6 小时间隔风暴轨迹的方法。该模型成功预测了未来 6 小时的风暴路径，准确率为 0.783，均方误差 (MSE) 为 0.0086。与 NOAA 方法和 GPRA 基准相比，我们的模型表现出更优的性能，并且训练速度也更快。

## 参考文献

- [1] "National Hurricane Center," National Oceanic and Atmospheric Administration, [Online]. Available: <https://www.nhc.noaa.gov/data/#hurdat>.
- [2] R. Chen, W. Zhang, and X. Wang, "Machine Learning in Tropical Cyclone Forecast Modeling: A Review," *Atmosphere*, vol. 11, no. 7, 2020.
- [3] H. Wang et al., "A Statistical Forecast Model for Atlantic Seasonal Hurricane Activity Based on the NCEP Dynamical Seasonal Forecast," *J. Climate*, vol. 22, no. 17, pp. 4481 – 4500, 2009.
- [4] NOAA, "NOAA Global Forecast System (GFS)," [Online]. Available: <https://registry.opendata.aws/noaa-gfs-bdp-pds/>.
- [5] NASA, "GOES Satellite Network," [Online]. Available: <https://www.nasa.gov/content/goes>.

- [6] S. Alemany et al., "Predicting Hurricane Trajectories using a Recurrent Neural Network," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 468-475, 2018.
- [7] T. Kim et al., "Near real-time hurricane rainfall forecasting using convolutional neural network models with Integrated Multi-satellite Retrievals for GPM (IMERG) product," *Atmospheric Research*, vol. 270, 2022.
- [8] R. Cahuantzi et al., "A comparison of LSTM and GRU networks for learning symbolic sequences," *arXiv:2107.02248*, 2021.
- [9] R. Bose et al., "A real-time prediction methodology for hurricane evolution using LSTM recurrent neural networks," *Neural Comput & Applic*, vol. 34, pp. 17491 – 17505, 2022.
- [10] A. Vaswani et al., "Attention Is All You Need," *Adv. Neural Inf. Process. Syst.*, 2017.
- [11] "Transformer' s Encoder-Decoder," [Online]. Available: <https://kikaben.com/transformers-encoder-decoder/>.
- [12] A. C. Amanambu et al., "Hydrological Drought Forecasting Using a Deep Transformer Model," *MDPI Water*, vol. 14, issue 22.
- [13] W. Jiang et al., "Transformer-based tropical cyclone track and intensity forecasting," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 238, July 2023.
- [14] NOAA, "National Hurricane Center Forecast Verification," [Online]. Available: <https://www.nhc.noaa.gov/verification/verify8.shtml>.
- [15] G. E. Birchfield, "Numerical Prediction of Hurricane Movement with the Use of a Fine Grid," *J. Atmos. Sci.*, vol. 17, no. 4, pp. 406 – 414, 1960.
- [16] "Formula to Find Bearing or Heading Angle Between Two Points: Latitude Longitude," [Online]. Available: <https://www.igismap.com/formula-to-find-bearing-or-heading-angle-between-two-points-latitude-longitude/>.
- [17] D. Hendrycks et al., "Gaussian Error Linear Units (GELUs)," *arXiv preprint arXiv:1606.08415*, 2016.