

硕士学位论文

基于深度学习的台风路径与 强度预测方法研究

RESEARCH ON TYPHOON MOVEMENT AND INTENSITY FORECASTING BASED ON DEEP LEARNING

徐光宁

哈尔滨工业大学

2020 年 6 月

国内图书分类号：TP183

国际图书分类号：004.8

学校代码：10213

密级：公开

工学硕士学位论文

基于深度学习的台风路径与 强度预测方法研究

硕 士 研 究 生：徐光宁

导 师：叶允明教授

申 请 学 位：工学硕士

学 科：计算机科学与技术

所 在 单 位：哈尔滨工业大学（深圳）

答 辩 日 期：2020 年 6 月

授予学位单位：哈尔滨工业大学

Classified Index: TP183

U.D.C: 004.8

A dissertation submitted in partial fulfillment of
the requirements for the academic degree of
Master of Engineering

**RESEARCH ON TYPHOON MOVEMENT AND
INTENSITY FORECASTING BASED ON DEEP
LEARNING**

Candidate:	Guangning Xu
Supervisor:	Prof. Yunming Ye
Academic Degree Applied for:	Master of Engineering
Speciality:	Computer Science and Technology
Affiliation:	Harbin Institute of Technology, Shenzhen
Date of Defence:	June, 2020
Degree-Confering-Institution:	Harbin Institute of Technology

摘要

我国地处世界台风高发的西北太平洋区域,每年夏秋台风季我国沿海地区都会因台风灾害造成不同程度经济损失和人员伤亡,准确预测台风的路径与强度对我国沿海地区的气象灾害防治有着重要的意义。目前基于深度学习的台风预测方法还处于探索阶段,已有研究大部分仅基于台风的二维时序特征,并没有充分利用台风的三维时序特征,从而无法融合不同模态之间的特征信息,易造成预测精度不高的问题。

本文重点针对台风预报中的多模态异构时序特征融合与挖掘问题,研究基于深度学习的台风路径预测和台风强度预测方法,主要内容如下:

(1) 在台风路径预测方面,针对预处理阶段的台风数据集质量评估问题,本文提出使用自相关系数评估数据集 CLIPER 特性的方法;针对模型训练阶段影响台风路径等压面的选择问题,本文引入残差学习的通道注意力机制(Residual Channel Attention, ReCA)以自动选择高响应等压平面;针对模型训练阶段的传统 ConvLSTM 无法提取大尺度深层非线性特征的问题,本文提出了迭代卷积 GRU 方法(Multi-ConvGRU),并在 wide&deep 框架下融合台风二维、三维时序异构特征。实验表明利用本文所提方法可有效融合多模态台风时序特征,较大提升了 24 小时台风路径预测精度。

(2) 在台风强度预测方面,针对模型训练阶段的台风强度风速分量学习问题,本文提出了一种基于空间注意力机制的融合网络(Spatial Attention Based Fusing Network, SAF-Net)模型。该模型将台风强度分解为 U Wind 变量和 V Wind 变量,并对每一个等压面引入空间注意力机制以提取风速分量的高响应区域,然后通过卷积融合对两种风速分量进行融合,最终实现通过学习各风速分量来预测合速度的目标。实验表明,使用本文提出的 SAF-Net 模型可提高 24 小时台风强度预测的精度。

基于台风路径与强度的研究成果,设计实现了一个基于 Flask 框架的台风路径与强度可视化系统,通过 Bootstrap 框架实现前端响应式布局、利用异步通信机制完成前端页面联动和后端数据更新,实现了台风路径和强度预报结果的可视化。

关键词: 台风路径预测; 台风强度预测; 深度学习; 时空序列学习

Abstract

China is located in the Western North Pacific basin where is the most active typhoon region in the world. Every year during the summer and autumn typhoon seasons, China's coastal areas will be suffered varying degrees of economic loss and casualties due to typhoons attack. Predicting the path and intensity of typhoons accurately is very important for the prevention and control of meteorological disasters in China's coastal areas. At present, statistical typhoon prediction based on deep learning methods is still in the exploration stage. Most of the current research is almost based on the 2D characteristics of the typhoon and does not consider well of the 3D characteristics of the typhoon. Additionally, the fusion of information of 2D and 3D typhoon model is lacking, which is likely to cause low prediction accuracy to both tasks.

This article focuses on the fusion and mining of multi-modal heterogeneous time-series features in typhoon forecasting and studies the methods of typhoon path prediction and typhoon intensity prediction based on deep learning. The main research works and results include the following three aspects.

In terms of typhoon path prediction, for the quality evaluation of the typhoon dataset in the preprocessing stage, I propose a method for evaluating the CLIPER characteristics of the dataset using Autocorrelation Coefficients. For the problem of selecting the isobaric planes of the typhoon path during the model training stage, I introduce Residual Channel Attention (ReCA) to automatically select high-response isobaric planes. In the model training phase, to solve the problem that traditional ConvLSTM cannot extract large-scale deep nonlinear features, I propose an iterative convolution GRU method (Multi -ConvGRU). Finally, I integrate the two-dimensional and three-dimensional chronological heterogeneous features of typhoon under the wide&deep framework. Experiments show that the method proposed in this paper can effectively integrate the multi-modal typhoon timing features, which greatly improves the 24-hour typhoon path prediction accuracy.

In terms of typhoon intensity prediction, for the problem of typhoon intensity component learning at the model training stage, I propose a Spatial Attention Based Fusing Network (SAF-Net) model based on spatial attention mechanism. The model decomposes the typhoon intensity into U Wind variable and V Wind variable, and introduces a spatial attention mechanism to each isostatic surface to extract the high-response region of the wind speed component, and then fuses the two wind speed

components through convolution fusion. Achieve the goal of predicting the closing speed by learning each wind speed component. Experiments show that using the SAF-Net model proposed in this paper can improve the accuracy of 24-hour typhoon intensity prediction.

Also, we designed and implemented a typhoon path and intensity visualization system based on the Flask framework. The front-end responsive layout was implemented through the Bootstrap framework, and the front-end page linkage and back-end updates were completed using asynchronous communication mechanisms. The system can show a typhoon path and intensity conveniently.

Keywords: typhoon path prediction, typhoon intensity prediction, deep learning, spatiotemporal sequence learning

目 录

摘 要.....	I
Abstract.....	I
第 1 章 绪论.....	1
1.1 研究背景和意义	1
1.2 国内外相关研究概况	2
1.2.1 基于动力学台风预报方法研究现状	2
1.2.2 基于统计学台风预报方法研究现状	4
1.2.3 基于统计动力学预报方法研究现状	6
1.3 问题的总结与分析	7
1.4 本文主要工作	8
1.5 本文组织架构	9
第 2 章 台风路径与强度预测相关背景知识	10
2.1 台风相关背景知识	10
2.2 台风路径与强度预测相关方法	12
2.2.1 基于最佳路径数据集的路径与强度预测方法	12
2.1.2 基于再分析数据集的路径与强度预测方法	14
2.3 相关评测指标	16
2.3.1 台风路径预测相关指标.....	16
2.3.2 台风强度预测相关指标.....	16
2.4 本章小结	16
第 3 章 基于多模态时空卷积的台风路径预测	17
3.1 引言	17
3.2 基于 CLIPER 的台风路径二维时序表征算法.....	18
3.2.1 台风 CLIPER 特性判定及马尔可夫性假设.....	18
3.2.2 基于 CLIPER 特性的台风路径二维时序表征.....	19
3.3 基于深度学习的台风三维时序表征算法	21
3.3.1 利用大气再分析数据构建台风的三维时序结构	21
3.3.2 台风三维时序表征的 ReCA 机制	22
3.3.3 台风三维时序表征的 Multi-ConvGRU 原理	25
3.4 基于 wide&deep 异构模态台风路径特征融合模型	26

3.5 实验结果与分析	29
3.5.1 实验环境	29
3.5.2 实验设计	29
3.5.3 实验数据	29
3.5.4 实验结果与分析	30
3.6 本章总结	33
第 4 章 基于多模态时空表示学习的台风强度预测	34
4.1 引言	34
4.2 基于 CLIPER 的台风强度二维时序表征算法	34
4.3 基于 SAF-Net 的台风强度三维时序表征算法	36
4.3.1 SAF-Net 中的空间注意力机制	36
4.3.2 SAF-Net 中的融合机制	39
4.4 基于 wide&deep 异构模态台风强度特征融合模型	41
4.5 实验结果与分析	42
4.5.1 实验环境	42
4.5.2 实验设计	42
4.5.3 实验数据	43
4.5.4 实验结果与分析	43
4.6 本章总结	46
第 5 章 基于 Flask 的台风路径与强度可视化系统	47
5.1 引言	47
5.2 系统架构设计	47
5.3 历史台风路径与强度可视化模块	48
5.4 预测台风路径与强度可视化模块	48
5.5 模型对比可视化模块	49
5.6 数据存储模块	49
5.7 人机交互界面	50
5.8 本章小结	54
结 论	55
参考文献	56
哈尔滨工业大学学位论文原创性声明和使用权限	62
致 谢	63

第 1 章 绪论

1.1 研究背景和意义

随着人类社会不断进步和科技不断发展，人类在不断认识世界、改造世界的社会实践过程中，也不断受到自然世界的规律所制约。在制约人类发展的自然规律中，自然灾害对人类的影响尤为显著，如地震、森林大火、火山喷发、台风等自然灾害都大范围影响着人类的活动与发展。为了更好地实现人类的可持续发展、减少自然灾害对人类经济社会的影响，科学家对自然灾害规律的探索从未间断，特别如我国科学家对台风灾害的探索更显突出。由于我国具有绵长的海岸线和地处台风高发的西北太平洋区域，每年夏秋季节的台风高发期，因台风登陆我国沿海地区所造成的经济损失和人员伤亡的报道屡见不鲜。虽然如此，但早期的统计台风预测和灾害防治工作由于受到科学技术发展的限制，其预测效果无法达到减灾避险的要求。

在深度学习、大数据、物联网等信息技术不断更迭和发展的条件下，分析客观复杂的事物技术愈发成熟。例如，从单一模态发展到多模态的数据采集技术大大增加了复杂事物的建模水平，这里所谓的模态是指某种事物发生或经历的方式，当研究问题包含同一事物发生的不同侧面时，研究问题便被定性为多模态的^[1]。同时，在数据愈发复杂的表征演化进程中，传统的时间序列也发展成为时间、空间相关联的时空序列，数据从传统的两个维度发展到目前涵盖空间、时间的四维结构。在多模态和时空序列的发展下，催生了多模态时空序列表示学习技术，而该技术则可以很好地应用在台风预测的研究中。这是因为台风的数据是多模态的，其主要通过多普勒天气雷达、飞机探空、卫星云图等方式采集其不同模态的数据。其次，台风的数据是时空化的，其包括经度、纬度、等压高度和时间四个维度，可以被表示为复杂的时空序列数据。

综上所述，由于深度学习能够更好地表征数据，且利用深度学习技术研究统计台风预测技术的条件已经具备。加之，准确预测台风的行进路径和中心强度对我国的自然灾害防治工作有着不可忽视的指导作用。因此，本课题的研究有着极其重要的现实意义。

1.2 国内外相关研究概况

台风路径与强度的预测问题关系到国家和地区的国计民生，因此国内外许多学者对其预测的方法和模型做了仔细深入的研究。台风的路径预测侧重于对台风及其周边大气环流背景的研究，环流的方向很大程度上决定台风的移动方向，因此属于大尺度问题。而台风的强度则侧重于台风的内在热动力结构与环境场交叠，台风内在热动力的分布及风力因素很大程度影响着台风的强度，因此属于多尺度问题。

虽然台风路径与强度两个任务的发生原理与问题性质不相同，但其建模思路和预测方法均相似。主要预测方法可以分为经验方法和客观方法两大类。其中，经验方法属于主观定量方法，往往通过天气图等气象因子并结合历史相似气象情况，预测未来的台风路径与强度^[2]。客观预报主要分为三种方法：基于动力学台风预报方法、基于统计学台风预报方法和基于统计动力学台风预报方法。目前，实务性的台风预测方法主要使用主客观方法结合得到最终预报结果。本文将主要介绍台风预测的客观方法的国内外研究现状，重点介绍基于统计学的台风预报方法。

1.2.1 基于动力学台风预报方法研究现状

基于动力学台风预报方法，又称为数值台风预报方法，是目前台风客观预报所采用的主要方法。该预测方法是基于空气动力学和热动力学原理，通过分析影响台风路径与强度的气象因子和物理过程如辐射、感热交换、地形影响等因素，构建大气动力学方程，以数学建模的方式预测未来台风的路径与强度。由于该方法所建立方程复杂难以求出解析解，因此往往需要借助大型、甚至超级计算机求解出其数值解以完成预测。动力学方法最早由学者V.Bjerknes^[3]于20世纪初提出并于1956年成书出版，该书根据某时刻实测的气象资料，通过描述大气运动的动力方程组，计算出未来某时刻大气的运动和状态，根据上述方法该书将天气预报问题转化为数学初值问题。该书提出的大气运动方程由五条偏微分方程组成包括：x方向运动方程、y方向运动方程、z方向运动方程、连续方程和热通量方程。得到大气运动方程后，再给出气象观测资料作为方程初值便可以利用数值分析求出方程的数值解。在V.Bjerknes提出利用数学的偏微分方程处理天气预测的问题后，数值预报方法有了很大的发展，逐渐发展成为了区域模式、全球模式和集合预报三大方向。

区域模式是基于某个特定区域的中尺度天气预报模式，不同的区域模式反映了特定区域的气象特点。比较有代表性的区域模式有MM5模式，该模式由Reisner JR^[4]等学者研究，由宾夕法尼亚州立大学和美国国家大气研究中心共同维护。MM5是一个有限区域模式，遵循地形的sigma坐标模型，用于复制或预测中尺度和区域尺度的大气环流，该模式自开发以来多次更新迭代以用于气象研究和天气预报业务。基于MM5模式，我国上海气象研究院顾建峰^[5]等学者探索出了一个适合预测华东地区气象预报的第五代区域预报模式，该模式以MM5为动力框架，采用三维最优插值方法对探空、地面等常规资料以及卫星、测厚、卫星测风报等非常规资料进行处理。在预报模式处理上，应用静力平衡方案，以非等距的形式将全球地形资料分为15层，最顶层为100pha，并截取 $95^{\circ} \times 95^{\circ}$ 的经纬度网格点，在以MM5的物理过程为参考选择出适用于华东地区的有关物理过程。此外，由沈学顺等^[6]学者研发的我国新一代数值预报系统，GRAPES (Global/Regional Assimilation and Prediction System)。GRAPES是完全由我国科学家研发，其内容包括：资料同化、模式动力框架、物理过程等数值预报基础业务于一身的大型软件系统，其中资料同化技术采用了变分同化技术并设计出3DVAR三维变化框架，该技术可以更加直接地同化如卫星、雷达等非常规资料，在GRAPES模式动力框架方面，采用全可压的完全方程组，通过该方程组可以提供全球/区域模式的一体化融合、三维矢量离散化、准单调正定水汽平流计算方案等功能。

全球模式是基于整个地球的气象系统的大尺度天气预报模式，该模式将全球大气系统看作是一个统一的整体，在全局层面考虑天气系统的物理运动情况，其缺点是粒度粗、精度不高。早期比较有代表性的是美国国家环境预报中心(National Centers for Environmental Prediction, NCEP)开发的GFS(Global Forecast System)，该全球模式可以处理十种大气和土地土壤变量，从温度，风和降水到土壤湿度和大气臭氧浓度。GFS覆盖了整个地球，网格点之间的基本水平分辨率为28公里。GFS是一个耦合模型，由四个单独的模型(大气模型，海洋模型，陆地/土壤模型和海冰模型)组成，它们一起工作以提供准确的天气状况图^[7]。沈学顺等^[8]学者在自主研发GRAPES模式系列后，拓展出了GRAPES_GFS(GRAPES Global Forecast System)全球数值预报系统并将其业务化，改进了原GRAPES模式多方面问题。在动力框架方面，改进了温垂直平流算法、极区滤波方案等；在物理计算方面，改进了积云对流、边界层等；在全球三维同化方面，提高了资料质量控制和误差订正技术，并提升了卫星资料同化技术。

集合预报的提出是由于大气系统是一个高度非线性的系统，任何的初始值都仅仅是真实大气运动的一个尽可能近似，而气象分析数据中存在的微小误差对最后大气预测的结果影响极大，因此将代表观测当前大气状态的不同初始值经过相同大气方程得到的预报结果的方法被称为集合预报^[9]。目前比较著名的集合预报系统(Global Ensemble Forecast System, GEFS)包括，ECWMF-GEFS，该系统采用了 50 个成员集合预测，初始不确定性方法采用了 Buizza 等^[10]学者提出的集合资料同化奇异向量法，该初始化方法主要通过对扰动观测、海表温度以及物理过程中的参数化倾向得到，拥有误差增长慢，集合离散程度大的优点，可以更好地追踪初始化时不确定性带来的时空演化，提供更好的概率分布以提高预报准确率。另一个比较著名的则是 NCEP-GEFS，该系统在初始场扰动方法中采用了重新尺度化集合转换技术来生成集合预报成员，该方法由 Bishop 等学者^[11]提出，其核心思想是对生成扰动的预报成员以接近正交的方式分解并保持依流型变化，该变化代表着误差最快增长的非线性扰动。此外，NCEP-GEFS 采用集合预报成员为 20 个，在尽量保证预报精度的前提下大大降低了计算资源。另外，NCEP-GEFS 还采用了涡旋重定位技术，即确定初始场涡度位置后将其从原有的环境场中分离出来迁移到观测位置，引入人为涡度位置在一定程度上可以加强初始场涡度观测比实际观察弱的缺点，可以达到减少预报发散度的效果。

1.2.2 基于统计学台风预报方法研究现状

基于统计学台风预报方法，是根据数理统计方法对大量的历史资料进行统计分析，找出天气变化蕴含的统计规律及预报因子与预报量之间的方程关系，一般采用回归方程建立统计数学模型。与动力学预报方式相比，统计学预报方法并不涉及大量复杂的数值计算，因此一定程度上节省了计算资源，但也存在无法覆盖统计以外的异常情况等问题。一般而言，统计学预报方法分为两个步骤即：筛选预报因子、建立预报方程。预报因子选择需要根据气象观测资料且优先选择物理意义上与预报量相关的预报因子，预报特征选择完后通常需要进行统计检定判断因子的统计特点。建立预报方程一般采用回归的方式，将所选择的预报因子与预报量建立一个多元线性回归方程。

早期预报方法主要以数理统计为主，其代表的方法为气候持续法(CLImatology and PERsistence, CLIPER)，该方法重点在于挖掘有效线性预报因子。CLIPER 于 1972 年由 Neumann^[12]提出，该模型提出是为了替换当时美国国家飓风中心的 HURAN 模拟系统。CLIPER 根据 HURAN 系统中的特征通

过逐步回归的方法选择出显著的预报因子进行模型拟合。不仅如此，论文对显著性不强的特征进行了细致分析，同时还分析了模型的平均误差。作者总结相比于 HURAN 模拟系统，CLIPER 具有计算简单、模型稳定、误差友好等特点。正因为这些优点，使得 CLIPER 成为台风路径与强度预测方法的基线模型。CLIPER 方法提出后，不少学者在数理统计思路下继续展开研究。其中比较代表性的是 Aberson^[13]的研究，该工作在 Neumann 的 CLIPER 基础上做了修改，相比于直接预测从开始时刻到未来的经度纬度，作者预测的是未来以 12 小时为周期的经度纬度，这样做可以更好地保证预测的连续性和合理性。此外，作者不仅对原始数据进行筛选以提升数据质量，更加合理地对预报因子进行筛选，还化简回归方程的推导过程，最终结果在误差与 Neumann 相接近的情况下，得到了比 Neumann 工作更好的预报效果。该工作不仅提升了 Neumann 的预报效果，而且得出 CLIPER 方法对大西洋区域的台风预测具有良好效果的结论，使得 CLIPER 被其他海洋区域借鉴成为可能。我国在借鉴该工作上，廖木星^[14]等学者做出了比较早期的探索，该论文结合我国西北太平洋区域 40 年的台风观测数据总结出我国台风运动规律，设计了适合我国海域的预报因子并通过逐步回归进行因子筛选后得到了经度、纬度的回归方程，并将该方程编写成业务系统，供西北太平洋地区船舶航行使用。该工作为远洋船舶、舰艇或岸上安监部门作台风路径预报及防御决策提供有效作用。

20 世纪末到 21 世纪初，越来越多研究者根据大气系统具有高维度非线性特性展开工作，因此研究重心从线性特点转向非线性特点，机器学习的方法开始广泛应用在台风预报问题上。比较有代表性的模型如支持向量机(Support Vector Machine, SVM)和人工神经网络(Artificial Neural Networks, ANN)。线性 SVM 模型是一种监督式的二分类模型，其原理是将样本表示为空间中的实例点，通过映射得到尽可能宽的间隔将两种不同的类别分开。非线性 SVM 采用了核技巧将其输入隐式映射到高维特征空间中再进行分类，因此非线性 SVM 比较好解决了高维数据分类问题。Song 等^[15]在分析当时数值预测台风技术的轨迹与真实轨迹偏离较远的立足点上，对输入数据进行降维后，利用 SVM 模型的核方法代替数值预测的非线性特征提取手段。利用机器学习的非线性特征提取能力来模拟台风的非线性特性的尝试，为后续台风路径的研究打开了一个新的思路。ANN 是一种具有很好的非线性统计性数据建模工具，其通过一个基于统计学的学习方法得以优化。Wang 等^[16]尝试使用 ANN 预测台风路径，对西北太平洋地区的台风数据进行最大最小归一化后输入到 ANN，与 CLIPER 和 SVM 方法不一样的是该输入仅包含了当前经纬度和前 6 小时经纬度，与之

前手动构建特征不同，该工作尝试利用 ANN 的非线性特点自动构建适合的预报因子。虽然该工作的预测结果并不理想，但却突破了长期手动构建预报因子的限制，尝试利用非线性自动构建特征，为未来研究的进一步发展开辟了一个新的思路。

当前，随着深度学习的迅猛发展，利用深度学习的强大非线性特征提取能力处理高度非线性的大气系统的研究路线逐渐走热。Sophie^[17]等学者利用 CNN 处理再分析数据，并将所提取的非线性特征与台风相关的经度、纬度、风速、气压、儒略日等特征融合，最后预测台风的经纬度。该工作虽然只是预测了东太平洋地区和大西洋地区的台风路径，但是其提出的将台风二维的路径特征与台风三维的大气特征，利用深度学习进行融合的想法突破了传统方法粗暴地将大气再分析数据进行降维的限制，更好地保留了台风的三维特性。该工作还尝试将不同的大气再分析数据拼接后以卷积的方式进行融合，并利用卷积神经网络的通道处理台风的时间维度，该文章的很多工作对本文的研究都有着重要的启发意义。Alemany 等^[17]学者利用 LSTM 为台风路径建立预测模型。与传统的预测直接预测经纬度不同，该工作将大西洋区域划分为了 $1^\circ \times 1^\circ$ 的经纬度网格点并对之进行编号，将台风每段路径中心经纬度归类到相应的格点中，该工作实际上通过编码的方式将经度、纬度编码到一个固定的数字，利用 LSTM 预测该编码，文章指出该方式可以有效减少直接预测带来的递归误差传递。此外，该工作还通过构造台风的位移特征加入到 LSTM 中，提高预测效果。作者的工作一方面通过划分网格的方式为台风路径预测模型提供了一个崭新的思路；另一方面，也为使用 RNN 预测台风时序问题做了很好的尝试。Kim 等学者^[19]根据 Shi 等^[20]学者提出的 ConvLSTM 结合大气再分析数据做了一个台风识别任务，通过 ConvLSTM 的时空维度交叠的学习，比仅使用 CNN 得到的效果要更好，该文章成功将 ConvLSTM 的模块应用到大尺度的时空数据处理，虽然仅使用了台风的三维特征并没有做模态融合，但却证明了 ConvLSTM 在时空关系非线性特征提取上比 CNN 更优秀。该工作首次将台风的三维特征以时空序列的方式进行处理，也为本文的任务提供了很好的参考价值。

1.2.3 基于统计动力学预报方法研究现状

基于统计动力学预报方法，是将动力学与统计学相结合的一种预报方法，其原理是利用数值预报的方法求解大气运动方程，得到方程的数值解后作出数值预报，最后再利用统计方法对预报的结果进行加工得出局地天气预报的结果。统计动力学预报方法一般可分为两种技术：第一种为完全预报方法，该方法根

据历史观测的资料和统计方法得到预报方程，并将数值预报的输出结果代入到该预报方程，因此该方法的预报结果的提高完全依赖于数值预报；第二种为模式输出统计方法，该方法通过数值预报的资料选取预报因子并根据选择出的预报因子构建预报方程，该方法的好处是比较充分考虑数值预报资料的特点，因而能够有效避免误差，此方法并不完全依赖于数值预报的结果。

由以上分析可知，统计动力学预报的方法重点是选择出与预报量强相关的预报因子。洪梅等学者^[21]对 T106 数值预报产品 500 hPa 位势高度场序列并对该资料进行经验函数正交分解(Empirical Orthogonal Function, EOF)，得到的时空分解变量作为统计动力模型的输入，其后利用遗传算法进行全局、并行优化参数使得模型在中长期预报效果比 T106 数值预报产品要好。杨玉震等学者^[22]等对 NCEP 再分析资料、T106L19 模式产品和热带气旋历史观测资料三种资料进行了研究分析，在原来仅使用逐步回归筛选预报因子的基础上，引入了主成分分析(Principal Component Analysis, PCA)方法，通过 PCA 方法使模型可以考虑众多预报因子的相关性从而选择出预报因子的主成分，达到降维效果。采用该方法后预测效果比仅使用逐步回归方法的效果更好。相比于其他学者研究如何更好提取特征，Zhang 等学者^[23]试图通过加入大尺度的气候变化模式，如太平洋子午线模式、大西洋子午线模式和北大西洋海表异常温度作为统计动力学方法中的预报因子。该方法出发点是试图从大尺度的范围考虑影响台风的预报因子，实验证明在上述区域加入大尺度特征后，效果得到了一定的提高。

1.3 问题的总结与分析

综上所述，台风路径与强度预测的主要方法分为三种：基于动力学台风预报方法、基于统计学台风预报方法和基于统计动力学台风预报方法。该三种预测方法各有优劣：对于动力学台风预报方法，需要根据动力学和热动力学的知识为未来的大气状态建立热动力学偏微分方程，但该方程组不存在解析解，因此需要利用数值分析的方法求其数值解。由于早期计算资源不足，无法获得精度很高的解析解。虽然目前计算资源充足，但由于大气资料和方程的复杂性，该方程的数值求解过程仍需要借助超级计算机进行计算，因而使预测成本过大而不易推广，且由于气象观测资料可能存在较大的误差，使得数值计算过程中容易出现误差传播导致解完全偏离，因此该方法对资料质量、观测技术和算力要求很高。对于统计学台风预报方法，该方法通过对大量气象历史资料总结出气象系统的统计特征，采用特征筛选等方法提取出显著性的预报因子后，以回归的方式建立预报方程。虽然所需要的计算资源相比于动力学台风预报方法减

少很多，但其预报方程仍存在缺乏动力学解释的问题。同时，虽然深度学习已经被应用在台风的三维特征提取中，但是仍存在无法提取深层大气非线性特征、忽略从大尺度与多尺度角度考虑台风三维建模等问题。对于统计动力学台风预报方法，虽然融合了统计学和动力学，但无论是完全预报方法还是模式输出预报方法，均依赖于动力学，该方法最大的贡献是通过统计学的方法，减少动力学预报的资料的误差，提升动力学台风预报的准确性，但本质上仍需要耗费大量的计算资源。

鉴于目前国内外本课题的研究现状，针对本文研究的台风路径与强度预测问题，设计的模型更需要关注大尺度的空间信息和长范围的时间信息，并融合不同模态的信息。因此本文对基于LSTM的时空序列模型进行改进，解决大尺度空间信息特征提取不充分的问题，同时解决空间信息缺乏对高响应区域特别处理等问题；此外，本文还研究讨论了基于CNN模型的时空序列分析的空间信息融合和提取高响应区域问题，通过融合不同分量的信息提升模型的预测准确性。本文主要针对以上提到的两个问题进行改进。

1.4 本文主要工作

综上所述，针对目前在台风路径与强度预测方法存在的问题，结合目前深度学习技术，本文做了以下三个方面的工作：

针对台风路径预测问题，本文引入自相关系数评估台风数据集质量并对台风路径采用了多模态时空卷积的方法进行建模：分别构建了台风路径二维时序特征和台风路径三维时序特征，利用 wide&deep 学习框架融合台风的两种异构模态。在处理台风三维时序表征中，本文引入了 ReCA 机制解决表征台风路径行进方向等压面选择问题和提出了 Multi-ConvGRU 解决大尺度空间特征提取能力不足的问题。

针对台风强度预测问题，采用多模态时序特征表示学习的方法，利用 wide&deep 学习框架融合台风强度二维时序特征和台风强度三维时序特征。结合台风强度为风速合速度特点，本文将强度分解成为 U Wind 和 V Wind 两个分量。由此，本文提出 SAF-Net 结构，在 SAF-Net 中引入了 Spatial Attention 机制使模型学习风速在等压平面的高响应部分，同时也引入卷积融合机制使得两种分风速分量合并成最终台风中心风速。

针对台风路径与强度预测可视化问题，本文利用 Flask 轻量级框架设计了一个基于响应式 bootstrap 框架的响应式台风路径与强度可视化 web 系统。通

过该系统可视化历史台风路径与强度、预测台风路径与强度及提出模型与基线模型的对比效果。系统展示了本文的研究成果。

1.5 本文组织架构

第 1 章绪论：本章为台风路径与强度预测方法的综述部分。首先，介绍台风路径与强度的研究背景及研究意义。然后，对本课题所研究的主要内容做了国内外研究现状分析，主要对台风路径与强度预测的动力学预报方法、统计学预报方法和统计动力学方法做了介绍。最后，介绍本课题的重点研究内容。

第 2 章台风路径与强度相关背景知识：本章主要介绍台风相关的业务知识。首先，介绍了台风生成过程和结构特点；然后，介绍全球台风高发地区，并介绍西北太平洋区域的台风路径与强度类别；最后，介绍台风的强度分类及不同台风模态观测的技术，并介绍台风路径和强度预测的评价指标。

第 3 章基于多模态时空卷积的台风路径预测：本章主要介绍台风路径预测的主要算法。首先，介绍了利用自相关系数方法评估二维台风数据集质量。接着，介绍了台风二维路径的 CLIPER 方法。然后，介绍了 Residual Channel Attention(ReCA)机制及迭代卷积 GRU 方法(Multi-ConvGRU)。最后，设计三个实验分别评估了所使用 CMA 数据的质量、CLIPER 方法的有效性及融合再分析数据后使用 ReCA 机制与 Multi-ConvGRU 模型的有效性。

第 4 章基于多模态时空表示学习的台风强度预测：本章主要介绍台风强度预测的主要算法。首先，介绍了台风二维强度的 CLIPER 方法。然后，介绍 Spatial Attention Fusing Network(SAF-Net)的结构及设计原理。最后，设计三个实验分别评估了二维台风强度 CLIPER 方法的有效性、SAF-Net 两个模块的有效性及其融合再分析数据后的有效性。

第 5 章基于 Flask 的台风路径与强度可视化系统：本章主要介绍台风路径与强度可视化系统的设计实现。首先，介绍了台风路径与强度可视化系统的系统架构设计。然后，分别介绍了数据存储模块、历史台风路径与强度可视化模块、预测台风路径与强度可视化模块、模型对比模块和人机交互界面。最后，对系统的所有功能进行详细的讲解。

第 2 章 台风路径与强度预测相关背景知识

本课题研究内容是基于深度学习的台风路径与强度预测方法，主要使用了中国国家气象局(China Meteorological Administration, CMA)最佳路径数据集^[24]和欧洲中期天气预报中心(European Centre for Medium-Range Weather Forecasts, ECMWF)的再分析数据。本文将根据这两种数据预测西北太平洋区域台风未来 24 小时的路径与强度。为了更好地了解本课题研究内容，首先介绍台风路径预测和台风强度预测相关的背景知识。

2.1 台风相关背景知识

台风是一种起源于赤道附近热带洋面的低压气旋。夏秋季节赤道附近的热带洋面日照充分，致海水被日照辐射蒸发上升到空中，形成低压系统。在地球的地转偏向力和日照对气压的影响下，被蒸发到空气中的水蒸气在北半球呈逆时针旋转、在南半球呈顺时针旋转，形成空气涡旋。若日照持续充分，该空气涡旋便会在蒸发的水汽和地转偏向力持续影响下不断增大，当其涡旋中心风速达到一定条件，便形成了台风。

台风一旦形成，便有了自身结构。一般而言，台风中心低层气流呈显著的辐合状态，高层气流呈显著的辐散状态，从台风中心依次向外大致可以分为以下三个区域：台风眼区、眼墙区、螺旋雨带区。风眼区的特点是风力最小且天气晴朗；眼墙区是台风中天气最为恶劣，经常狂风呼啸，大雨如注；螺旋雨带区则会常常伴有阵雨和大风。具体结构如图 2-1 所示。

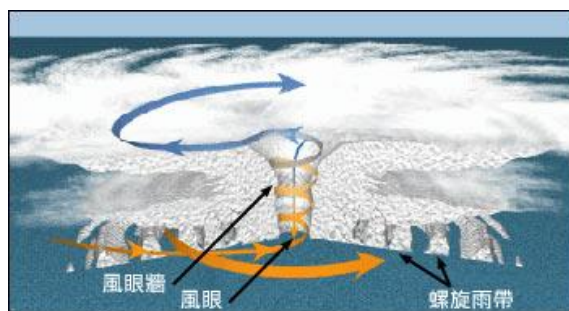


图 2-1 台风结构

台风的形成具有一定的区域特点。一般而言，全球台风高发区域一共分为七大区域，从序号 1 到序号 7 分别是：大西洋区域、东北太平洋区域、西北太

平洋区域、北印度洋区域、西南印度洋区域、东南印度洋区域和西南太平洋区域。具体如图 2-2 所示。

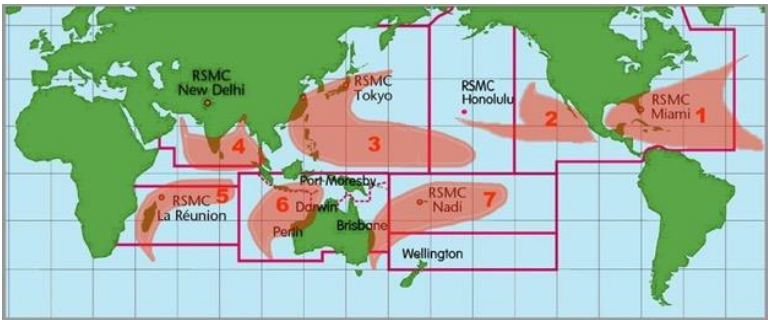


图 2-2 全球台风高发区域划分

由于我国毗邻西北太平洋区域，而该区域更是全球台风最高发的区域，因此研究西北太平洋区域台风路径的特点，对我国沿海地区灾害防控有重要意义。在西北太平洋区域的台风生成后，主要受到副热带高压的影响。由于副热带高压中的气压梯度指向赤道，在地球的地转偏向力的作用下，容易形成盛行西风。因此西北太平洋区域形成的台风一般具有西行特点。西北太平洋的台风路径主要分为：西进型、西北型、转向型和迷走型。西进型台风路径主要是受到盛行西风的影响，当副热带高压出现南移的情况，路径会出现偏南。西北型台风路径主要是盛行西风的北方向分量增加导致；转向型台风路径，主要是台风先西进转北后再往东北方向前进，形成一条抛物线形状的台风，出现该路径可能的原因是台风行进时穿过了副热带高压区域或其他地域因素导致；迷走型台风，可能会出现停滞、打转等情况，目前发生机理尚不明确，可能是受到不明气流引导或其他天气系统影响所导致的。

了解台风的路径可以使我们更好地预测台风登陆的自然规律，而了解台风的强度则可以让我们更好地了解台风登陆后所造成的破坏程度。台风的强度是指台风中心的风力大小，我国一般采用 2 分钟持续最大风速来判定台风的强度。根据台风中心风力的大小，我国对台风强度等级做出了如表 2-1 的划分。

表 2-1 台风强度等级划分

台风等级	风速数值
弱热带低压	≤10.8m/s
热带低压	10.8m/s - 17.1m/s
热带风暴	17.2m/s - 24.4m/s
强热带风暴	24.5m/s - 32.6m/s
台风	32.7m/s - 41.4m/s
强台风	41.5m/s - 50.9m/s
超强台风	≥51.0m/s

此外，更好的台风观测技术才能对台风的路径与强度等信息做出准确预测，从而减少台风对沿海国家和地区所造成的灾害。由于台风是一个庞大复杂的气候系统，单一的观测技术难以刻画台风全貌，目前主要有以下台风观测技术：气象遥感卫星观测技术、岛屿和油气平台站观测技术、船舶观测技术、锚定和漂流浮标站观测技术、多普勒天气雷达技术和飞机探空观测技术等。气象遥感卫星观测技术主要通过静止卫星的可见光，红外和水汽通道的云图来对台风进行观测、极轨卫星主要提供更高空间分辨率的海表温度、气溶胶光学厚度，进行大气温度和水汽轮廓线的红外和微波遥感；岛屿和油气平台站观测技术主要通过无人岛屿和油气平台对所在区域的海温、海盐、海浪等重要参数进行观测；船舶观测技术主要通过观测船对台风的风向、风速、最大风速、气压和温度等重要气象参数进行观测；锚定和漂流浮标站观测技术主要作用是在恶劣气候条件下完成海面气压、温度、海表温度和海浪等参数的观测；多普勒天气雷达技术主要是对台风带来的雷暴和定点降雨进行观测，雷达可以提供台风的风速和风向的分布，为天气预测提供动力学依据；飞机探空观测技术可以提供台风内部结构的云层、降水等参数，并且能实时获得台风内部的温度和风场分布等信息。

2.2 台风路径与强度预测相关方法

台风的路径和强度预测主要分为：短期预测(24 小时后)、近期预测(48 小时后)、中期预测(72 小时后)、中长期预测(96 小时后)和长期预测(120 小时后)，本文的任务是对台风中心的经纬度做短期预测，对台风中心的风速做短期预测。该两个任务所用的数据源相同，预测方法基本通用。

2.2.1 基于最佳路径数据集的路径与强度预测方法

(1) CLIPER 方法

CLIPER方法是台风路径与强度预测的最通用的方法，其本质是一个多元线性回归模型，该模型主要通过构建台风的气候性因子和持续性因子的特征，拟合CLIPER预测方程并预测未来的台风中心经纬度与风速。其表达式如式(2-1)所示：

$$y = \mathbf{w}^T \mathbf{x} + b \quad (2-1)$$

其中， y 表示模型预测值， $\mathbf{x} = [x_1, x_2, \dots, x_n]$ 表示气候性因子和持续性因子

n 为因子总数, $w = [w_1, w_2, \dots, w_n]$ 为回归模型参数, b 为模型的偏置。在(2-1)式中, 气候因子是指由再分析数据所刻画的大气环境中的各种影响气候变化的因子如: 风速、气温、湿度等等的信息。由于再分析数据维度高, 往往需要降维处理后才作为 CLIPER 方法的因子; 持续因子是指台风作为一个二维平面中曲线, 其轨迹具有一定的运动规律性。该部分特征一般根据台风的中心经度、中心纬度、中心风速和中心气压等基础描述信息, 以及常用物理与统计变量构建出的因子。

(2) RNN模型

深度学习是一种具有强大表征能力的方法, 随着深度学习技术的不断发展, 目前多种深度学习框架已经在多个领域取得突破性发展, 其中RNN便是其中一种重要的深度学习框架。RNN是一种以序列数据为输入, 并不断地在序列的演进方向进行递归, 且所有神经单元链式连接在一起的神经网络。自Hochreiter^[25]于1991年提出长短期记忆网络(Long Short-Term Memory networks, LSTM)解决RNN模型无法学习长时间序列信息的长期依赖问题后, RNN模型逐渐走向成熟。LSTM针针对长期依赖问题新增了状态单元 c , 该单元只受当前时刻输入和上一时刻的输出影响, 因此长期记忆信息得到改善。此外, 新增的遗忘门 f 控制了上一时刻输出和记忆单元对当前单元的影响程度, 输入门 i 则控制了当前输入对状态单元的影响程度, 输出门 h 则控制当前输出内容的输入比例。LSTM原理图如图2-3所示。

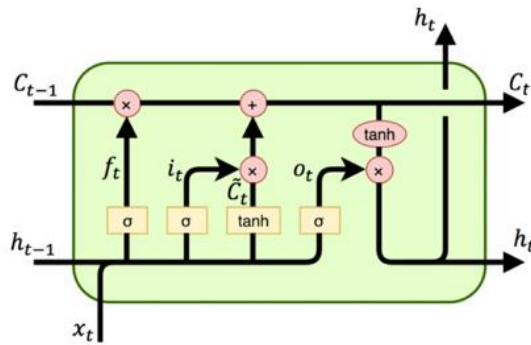


图2-3 LSTM原理图^[25]

其中, 遗忘门计算公式如(2-2)式所示:

$$f_t = \sigma(U_f x_t + W_f h_{t-1} + V_f c_{t-1}) \quad (2-2)$$

输入门计算公式如(2-3)式所示:

$$i_t = \sigma(U_i x_t + W_i h_{t-1} + V_i c_{t-1}) \quad (2-3)$$

当前记忆单元计算公式如(2-4)-(2-5)式所示:

$$\tilde{c}_t = \tanh(U_c x_t + W_c h_{t-1}) \quad (2-4)$$

$$c_t = i_t \odot \tilde{c}_t + f_t \odot c_{t-1} \quad (2-5)$$

输出门计算公式如(2-6)-(2-7)式所示：

$$o_t = \sigma(U_o x_t + W_o h_{t-1} + V_o c_{t-1}) \quad (2-6)$$

$$h_t = o_t \odot \tanh(c_t) \quad (2-7)$$

式中 \odot 表示哈达玛乘积。

2.2.2 基于再分析数据集的路径与强度预测方法

(1) CNN模型

卷积神经网络(Convolutional Neural Networks, CNN)是一种以卷积计算为核心的前馈神经网络，虽然CNN一般用于解决计算机视觉的问题。而大气再分析数据是描述地球三维空间大气环境的重要数据，其等压平面构成了地球的三维空间表征，若将这些等压平面看作是描述大气的特征图，则可以使用CNN模型对其进行处理。一般CNN分为2D CNN和3D CNN两种卷积神经网络。2D CNN最早于1989年由LeCun Yann^[26]提出，目前大多数研究都基于该结构，其具体模型如图2-4所示。LeNet 包含了深度学习的基本网络层：卷积层、池化层、全连接层等。其中卷积层是该网络的核心单元，该单元通过不同初始设置不同形状的卷积核，对输入的图像在不同位置做卷积操作，依靠监督学习的方式提取图像的特征。卷积操作具体公式表示如下：

$$H(i, j) = (X * W)(i, j) + b = \sum_{k=1}^n (X_k * W_k)(i, j) + b \quad (2-8)$$

其中， $H(i, j)$ 表示对应图像矩阵的第 i 行第 j 列的输出结果； $*$ 代表卷积操作； X, W 分别表示输入图像矩阵和卷积核； b 表示为偏置项； X_k, W_k 分别表示输入图像矩阵和卷积核矩阵的第 K 维图像矩阵和卷积核矩阵。

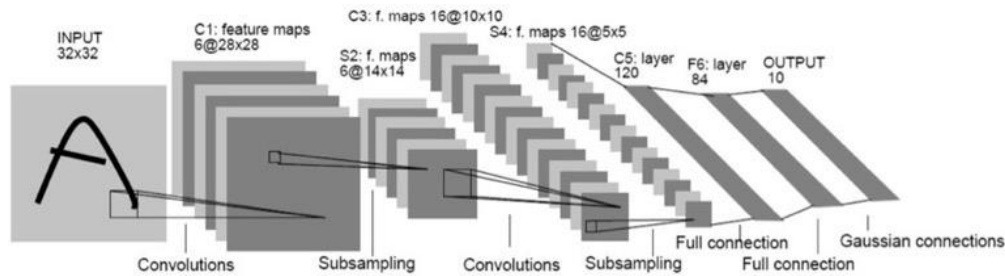


图2-4 LeNet原理图^[26]

虽然2D CNN可以很好地提取图像的二维特征，但是客观世界却是三维空间，2D CNN难以提取三维立体图像的特征。为了克服三维特征提取的问题Ji^[27]等人提出了3D CNN，该网络结构与2D CNN的不同在于：2D CNN的二维卷积

核之间相互独立并不相互关联，而3D CNN的三维卷积核由多个二维卷积核构成，且这些卷积核之间相互关联，共同提取空间上的垂直信息，达到三维立体卷积的效果，其原理图如图2-5所示。

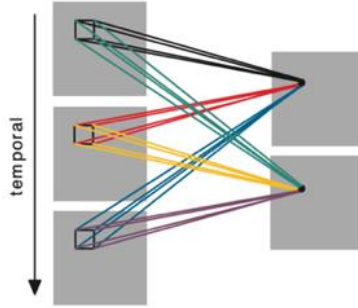


图2-5 三维卷积原理图^[27]

(2) ConvLSTM模型

LSTM在时间序列学习中取了良好的效果，但其在涉及时空问题上的效果却并不理想。虽然CNN可以通过采用2D CNN或3D CNN可以完成对二维或三维空间的表征，但却没有很好考虑时间维度上的空间信息表征。Shi等人^[20]在保证提取时间特征的基础上，在LSTM单元加入用于提取空间特征的卷积运算，提出了ConvLSTM模型并在时空数据学习中取得了非常好的效果。

与LSTM模型相比ConvLSTM将模型的所有输入、输出、隐含状态以及各个控制门都改成了三维张量，以方便模型通过卷积操作表征空间信息，而其余的计算方法与LSTM相同。ConvLSTM的内部结构如图2-5所示。

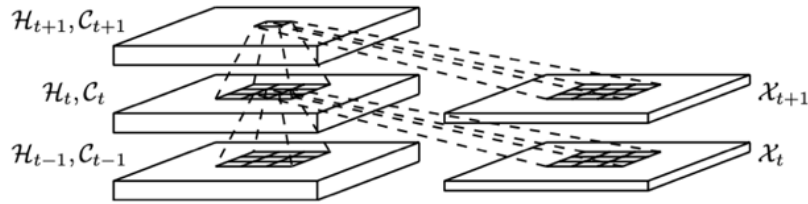


图2-5 ConvLSTM内部结构^[20]

其中，遗忘门计算公式如(2-9)式所示：

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f) \quad (2-9)$$

输入门计算公式如(2-10)式所示：

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i) \quad (2-10)$$

当前记忆单元计算公式如(2-11)式所示：

$$C_t = f_t \circ C_{t-1} + i_t \circ \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \quad (2-11)$$

输出门计算公式如(2-12)-(2-13)式所示

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} \circ C_t + b_o) \quad (2-12)$$

$$\mathcal{H}_t = o_t \circ \tanh(C_t) \quad (2-13)$$

式中*表示卷积运算； \circ 表示哈达玛乘积。

2.3 相关评测指标

在台风路径预测任务中，使用了球面上真实经纬度点与预测经纬度点的两点距离误差指标来评测预测结果。在台风强度预测任务中，通过比较预测风速与真实风速之间的绝对误差来评价预测结果。

2.3.1 台风路径预测相关指标

使用球面距离误差公式表示预测经纬度点与真实经纬度点之间的差异：

$$D = 2R \sin^{-1} \left(\sqrt{\sin^2 \left(\frac{|a_p - a_r|}{2} \right) + \cos a_p \cos a_r \sin^2 \left(\frac{|g_p - g_r|}{2} \right)} \right) \quad (2-14)$$

其中， R 表示地球半径一般取值为 6371.009km， a_p, a_r 分别表示预测纬度和真实纬度， g_p, g_r 分别表示预测经度和真实经度。

2.3.2 台风强度预测相关指标

使用均绝对方误差(Mean Absolute Error, MAE)来表示预测风速与真实风速之间差异：

$$MAE = \frac{1}{n} \sum_{i=0}^n |w_p - w_r| \quad (2-15)$$

其中， w_r 表示真实风速， w_p 表示预测风速。

2.4 本章小结

本章主要介绍了台风的背景知识和台风预测的相关基础知识。首先，介绍台风的生成原理、系统结构、路径特点和强度分类等基础知识。接着，介绍基于最佳路径数据集的 CLIPER 方法和 RNN 模型。然后，介绍基于再分析数据集的 CNN 模型与 ConvLSTM 模型。最后，对台风路径和台风强度的常用评价指标做了阐述。

第3章 基于多模态时空卷积的台风路径预测

3.1 引言

为减少台风对登陆地造成的人员伤亡和经济损失,需要对台风的路径进行准确的预测。良好预测结果的前提是合理地对历史台风数据进行建模,台风被非同源的数据刻画其多模态信息。在统计台风预测技术中,对台风仅仅使用某一个模态进行学习,难以达到准确预测的效果。由于不同模态之间的信息常常有互补的作用,因此融合不同模态之间的信息是提高统计台风预测效果的关键步骤。由于,短期台风路径预测任务的核心是预测 24 小时台风中心的经纬度,因此本章在多模态时空序列表示学习的基础上,以回归方式进行台风路径预测。

对现阶段台风路径预测模型进行分析,发现目前台风路径统计预测的研究工作大部分存在以下问题:目前对二维台风数据集的质量评估方法存在缺失且大部分研究仅仅关注二维台风路径^{[16][17]},目前对深层非线性空间特征的提取研究并不完备^[19],三维台风建模无充分考虑台风影响路径的等压面及台风三维结构^[17]。为了解决存在的这些问题,模型需要融合台风的不同的模态,需要提高深度非线性特征的提取能力,更需要模型从大尺度范围考虑影响台风行进的等压面。

对二维台风数据集的质量评估问题,采用自相关系数评估二维台风的 CLIPER 特性;对大部分研究仅仅关注二维台风路径的问题,本章提出了融合异构台风二维与三维模态;在台风三维特征学习方面,本文提出了时空卷积方法,该方法包括 Multi-ConvGRU 和基于残差的通道注意力(Residual Channel Attention, ReCA)机制,这两个方法分别解决了目前无法提取再分析数据的深层非线性特征的问题和没有考虑台风受等压平面影响及台风三维结构的大尺度问题。此外,本章采用 Google 提出的 wide&deep 联合学习框架对台风二维和三维两个异构模态进行融合。为了进一步评估模型的有效性,模型在 CMA 数据集与 ECWMF 的 ERA-interim 数据集上系统地进行了实验分析。实验结果表明:与现有方法相比,加入台风三维结构的模型能有效提升台风路径预测的准确性。

3.2 基于 CLIPER 的台风路径二维时序表征算法

3.2.1 台风 CLIPER 特性判定及马尔可夫性假设

气候持续性是指在气候的不断变化过程中,某些特殊气候状态在很长一段时间可以维持并反复出现的特点。因此 CLIPER 方法根据某种特殊气候出现的统计特性对其未来状态进行预测。

引入时间序列的自相关系数对 CLIPER 特性进行量化。对于一个时间序列 $X_{1,t} = \{X_1, X_2, X_3, \dots, X_t\}$, 若存在 k 阶滞后数及 k 阶向前数, 使得时间序列 $X_{1,t}$ 的 k 阶自相关系数为正相关, 则说明该时间序列具有 k 阶 CLIPER 特性; 其中, 某序列存在 k 阶滞后数是指序列存在正整数 k 使得当前时间步可以回溯 k 步, 即存在 $X_{1,t-k}$; 同理, 某序列存在 k 阶向前数是指序列存在正整数 k 使得开始时间步可以向前 k 步, 即存在 $X_{k,t}$; 图 3-1 举例说明了滞后数、向前数、向前序列和滞后序列的概念。

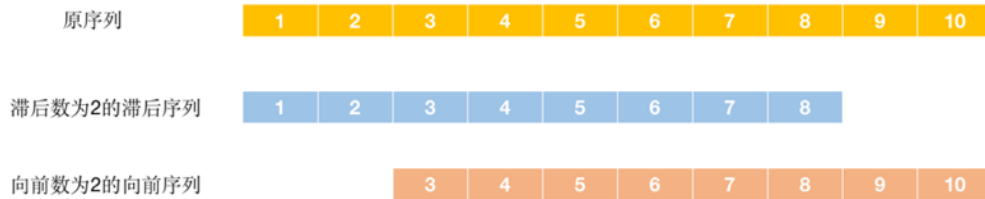


图 3-1 滞后、向前序列说明图

时间序列 $X_{1,t}$ 的自相关系数是度量同一个时间序列的 k 阶滞后序列和 k 阶向前序列的相关性, 自相关系数 R_k 如(3-1)式子所示:

$$R_k = \frac{\sum_{i=1}^{n-k} (X_i - \bar{X})(X_{i+k} - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (3-1)$$

其中, n 表示时间序列 $X_{1,t}$ 的时间步长; X_i 表示为时间序列 $X_{1,t}$ 的第 i 个时间步; \bar{X} 表示时间序列 $X_{1,t}$ 所有时间步的均值。

当台风路径具有 k 阶 CLIPER 特性的时候, 即说明该台风路径具有持续性的, 若 k 阶的 CLIPER 值越大, 则台风持续性越强, 未来台风路径受历史前 k 步路径影响更大。虽然天气系统具有非常大的复杂性和不确定性, 但是未来的天气发展一般只与当前或早前的大气系统状态有关, 因此, 利用 CLIPER 值度量台风的持续性不仅可以度量台风天气系统的反复性, 更能够度量其未来路径的马尔可夫性。为此, 在台风路径实际运动规律下, 我们不妨假设台风未来的路径仅与当前及前 k 步路径有关与其他时间步路径无关, 即台风路径仅具有 k 阶马尔可夫特性, 其数学表达如式子(3-2)所示:

$$Pr\{X_{n+1}|X_n, X_{n-1}, \dots, X_0\} = Pr\{X_{n+1}|X_n, X_{n-1}, \dots, X_{n-k}\} \quad (3-2)$$

其中， X_n 表示当前时间步， X_{n+1} 表示未来预测的时间步， X_{n-i} 表示历史时间步*i*。根据台风的气象背景知识，台风路径一般要求具有 4 阶 CLIPER 特性，因此本章假设未来台风路径仅具有 4 阶马尔可夫性。

3.2.2 基于 CLIPER 特性的台风路径二维时序表征

现根据 CMA 数据集的台风路径数据(见表 3-1)构建台风路径二维时序特征。

表 3-1 台风路径信息表

字段	记号	备注
台风编号	ID	台风国际编号“台风发生年份-当年台风序号”
台风发生时间	TIME	日期格式为：YYYY-MM-DD-HH，每 6 小时记录一次
台风中心经度	LAT	北纬为正，单位是0.1°N
台风中心纬度	LONG	东经为正，单位是0.1°E
台风中心风速	WND	台风中心两分钟持续最大风速，单位是米每秒，m/s
台风中心气压	PRES	单位是百帕斯卡，hPa

根据台风的 4 阶 CLIPER 特性、4 阶马尔可夫性及当前台风和前 24 小时的台风路径信息，构建台风路径二维时序特征表(见表 3-2)。为方便表达现作如下约定：LAT 表示当前时间台风中心纬度， LAT_i 表示前*i*小时台风中心的纬度， $LOC = (LAT, LONG)$ 表示为地球上的一点， $PATH_{i,j} = (LOC_i, LOC_j)$ 表示一段台风的路径；其中，LONG、WND、PRES、LOC、PATH 关于时间下标的表示和 LAT 的约定类似。

表 3-2 台风路径二维特征表

序号	特征	备注
1-5	LAT、 LAT_6 、 LAT_{12} 、 LAT_{18} 、 LAT_{24}	前 24h 纬度
6-10	LONG、 $LONG_6$ 、 $LONG_{12}$ 、 $LONG_{18}$ 、 $LONG_{24}$	前 24h 经度
11-15	WND、 WND_6 、 WND_{12} 、 WND_{18} 、 WND_{24}	前 24h 风速
16	Moth	台风发生的月份
17-18	$LAT - LAT_6$ 、 $LAT_6 - LAT_{12}$	前 6h-12h 一阶纬度差
19-20	$LAT_{12} - LAT_{18}$ 、 $LAT_{18} - LAT_{24}$	前 12h-24h 一阶纬度差
21-22	$LONG - LONG_6$ 、 $LONG_6 - LONG_{12}$	前 6h-12h 一阶经度差
23-24	$LONG_{12} - LONG_{18}$ 、 $LONG_{18} - LONG_{24}$	前 12h-24h 一阶经度差
25-26	$WND - WND_6$ 、 $WND_6 - WND_{12}$	前 6h-12h 一阶风速差
27-28	$WND_{12} - WND_{18}$ 、 $WND_{18} - WND_{24}$	前 12h-24h 一阶风速差

表 3-2(续表) 台风路径二维特征表

序号	特征	备注
29	$\sum_{i=0}^3 (\text{LAT}_{6i} - \text{LAT}_{6(i+1)})^2$	一阶纬度差平方和
30	$\sum_{i=0}^3 (\text{LONG}_{6i} - \text{LONG}_{6(i+1)})^2$	一阶经度差平方和
31	$\sqrt{\sum_{i=0}^3 (\text{LAT}_{6i} - \text{LAT}_{6(i+1)})^2}$	一阶纬度差平方和开根号
32	$\sqrt{\sum_{i=0}^3 (\text{LONG}_{6i} - \text{LONG}_{6(i+1)})^2}$	一阶经度差平方和开根号
33-34	$\sqrt{\text{LAT}}$ 、 $\sqrt{\text{LONG}}$	当前纬度、经度开根号
35-36	$\text{ACC}(\text{LOC}_6, \text{LOC})$ 、 $\text{ACC}(\text{LOC}_{12}, \text{LOC}_6)$	前 6h-12h 物理加速度
37-38	$\text{ACC}(\text{LOC}_{18}, \text{LOC}_{12})$ 、 $\text{ACC}(\text{LOC}_{24}, \text{LOC}_{18})$	前 12h-24h 物理加速度
39-40	$\text{Angle}(\text{LOC}, 0^\circ\text{N})$ 、 $\text{Angle}(\text{LOC}_6, 0^\circ\text{N})$	前 6h 位置与纬线夹角
41-42	$\text{Angle}(\text{LOC}_{12}, 0^\circ\text{N})$ 、 $\text{Angle}(\text{LOC}_{18}, 0^\circ\text{N})$	前 18h 位置与纬线夹角
43-44	$\text{Angle}(\text{LOC}, 0^\circ\text{E})$ 、 $\text{Angle}(\text{LOC}_6, 0^\circ\text{E})$	前 6h 位置与经线夹角
44-46	$\text{Angle}(\text{LOC}_{12}, 0^\circ\text{E})$ 、 $\text{Angle}(\text{LOC}_{18}, 0^\circ\text{E})$	前 18h 位置与经线夹角
47	$\text{Angle}(\text{PATH}_{0,6}, \text{PATH}_{6,12})$	前 6h-12h 路径夹角
48	$\text{Angle}(\text{PATH}_{6,12}, \text{PATH}_{12,18})$	前 12h-18h 路径夹角
49	$\text{Angle}(\text{PATH}_{12,18}, \text{PATH}_{18,24})$	前 18h-24h 路径夹角
50-51	$\text{Angle}(\text{LOC}, \text{LOC}_6)$ 、 $\text{Angle}(\text{LOC}_6, \text{LOC}_{12})$	前 6h-12h 台风移向角
52-53	$\text{Angle}(\text{LOC}_{12}, \text{LOC}_{18})$ 、 $\text{Angle}(\text{LOC}_{18}, \text{LOC}_{24})$	前 6h-24h 台风移向角

表 3-2 中台风路径二维特征包括了台风二维结构的基本表征参数，其中特征 1-32 为台风的线性信息，表示台风系统的线性结构：特征 1-15 是当前和历史台风基本信息，表征台风的当前和历史二维结构；特征 16 是台风发生月份，表征台风年度持续性的特点；特征 17-28 是当前和历史台风变化信息，表征台风的二维结构历史变化；特征 29-32 是当前和历史台风变化加速信息，表征台风的二维结构的加速情况；特征 33-57 为台风的非线性信息，表征台风系统的非线性结构：特征 33-38 对台风的经纬度做非线性变换、特征 39-57 对台风位置、路径信息做非线性变换。

3.3 基于深度学习的台风三维时序表征算法

3.3.1 利用大气再分析数据构建台风的三维时序结构

台风是一种具有三维立体结构的复杂、非线性的热带低压系统。仅仅使用台风的二维特征而不考虑其三维立体结构和台风所在的风带、暖冷流等气候等因素，无法对台风路径进行非常准确的建模。传统台风路径预测方法主要将台风视为二维平面的线段，而不考虑其三维立体结构；少部分研究工作虽然考虑了台风的三维结构，但却存在非线性空间特征提取不充分，没有考虑台风受等压平面影响的大尺度问题。为此，本小节将利用大气再分析数据构建台风及其周边的三维时序结构。

台风的行进路径主要受到两大因素影响：第一个是台风的内部动力因素，由于台风是一种强涡旋状的低压系统，其内部主要受到气压梯度力和科里奥利力影响，当其力分布产生变化的时候，台风的整体结构就会偏向于受力不均匀的一方移动；第二个是台风的外部环境因素，台风所在的环境流场直接决定着台风的移动方向。台风一般生成于低纬季风槽的辐合带，辐合带中的风场常处于对冲状态，极易形成切变涡旋，涡旋产生后台风的移动方向受对冲状态强的一方影响。同时，北半球副热带高压是个很大范围的顺时针环流，这些气流的方向作用于台风上，最终形成台风的行进路径。

基于以上两大因素，采用 ECMWF 的 EAR-Interim 数据集的重力势变量 (Geopotential) 构建台风及其周边的三维时序，该变量主要描述台风内部能量结构和动力情况是一种四维栅格数据其维度包括：经度、纬度、等压高度和时间。具体构建过程如下：

首先，将地球划成由单位为 $1^\circ \times 1^\circ$ 的经度、纬度曲面组成的球面，如图 3-5 所示。

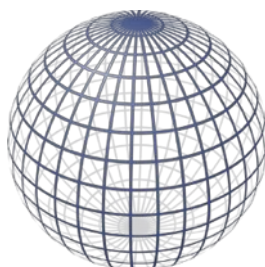


图 3-5 地球划分为 $1^\circ \times 1^\circ$ 网格点

然后，以台风中心经度、纬度为中心截取一个范围为 $31^\circ \times 31^\circ$ 的曲面，以囊括台风内部大气再分析信息与外部环境信息，如图 3-6 所示。

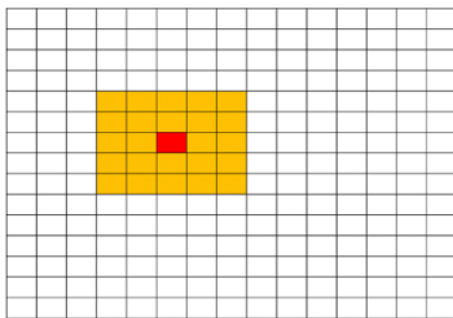


图 3-6 按台风中心划分 $31^\circ \times 31^\circ$ 网格点

最后，以台风为中心截取的二维平面，对不同等压面的采样构建台风及其周边大气环流三维结构，并在不同时间步完成以上步骤。如图 3-7 所示。

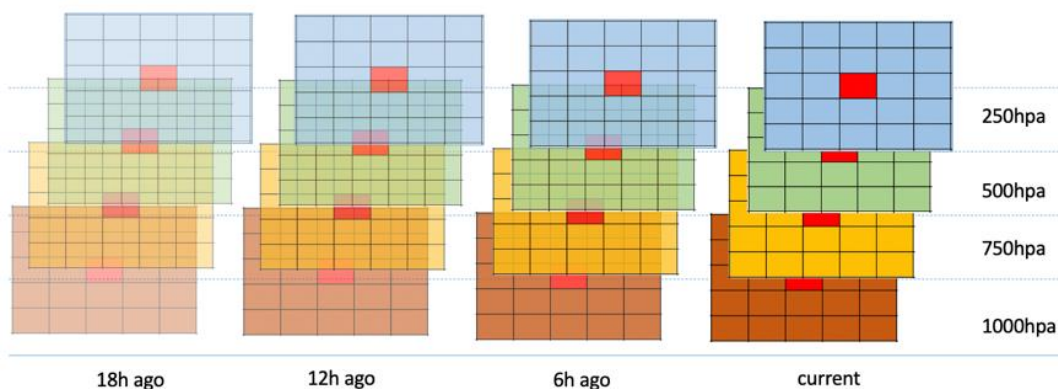


图 3-7 台风及环流三维时序结构

通过上述步骤，在经度、纬度、高度、时间四个维度上，根据大气再分析数据构建了台风及其周边的三维时序立体结构。

3.3.2 台风三维时序表征的 ReCA 机制

构建台风及其周边的三维时序结构后，学习该结构的规律是一个复杂的时空学习问题。本文引入了基于残差学习的通道注意力机制(Residual Channel Attention, ReCA)^[28]，以考虑影响台风前进的高响应等压面及台风的三维结构。

在深度学习的初期，研究者都认为网络层次的越深，伴随而来的性能效果越好。Alex^[29]等人设计出具有 7 层深度的 AlexNet 于 2012 年 ImageNet LSVRC 夺得冠军，并且引起了广大研究人员对深度学习的研究。随后，由牛津大学视觉组提出的 VGGNet^[30]在 2014 年 ImageNet 的定位比赛获得第一名，分类比赛获得第二名，其两个基础结构 VGG16 和 VGG19 在原有 AlexNet 做部分修改的前提下，进一步加深网络的层数并证明了增加网络层数的深度能够在一定程度上影响网络最终的性能。同年的 ImageNet 分类比赛的第一名则由具有 Inception 结构的 GoogLeNet^[31]获得，GoogLeNet 总共有 22 层网络结构，比

VGGNet 更深，但是由于采用了 Inception 结构，其参数比 VGGNet 更少。GoogLeNet 的成功更加标志着深度学习往网络层数更深的方向发展。

一般而言，神经网络的层数越深，其表达能力越强、越抽象和越能提取更深的非线性特征。但是网络层数并不能无限加深。随着网络的不断加深会出现梯度消失或梯度爆炸问题、过拟合问题和退化问题，这些问题将会导致网络性能的严重下降。2015 年，He 提出基于残差学习的 ResNet^[32]，一般网络深度不断加深会伴随着性能的急剧下降，而通过残差学习的方法在神经网络中加入跨层连接，可以将当前层的信息几乎不变地传到下一层。使得网络更关注于残差信息的学习，提高学习的效率。从而允许网络的层次不断加深提高网络性能。常见的 ResNet 有 50 层、101 层与 152 层。

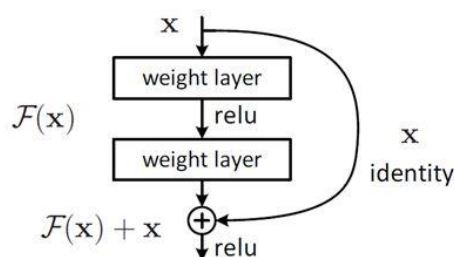


图 3-8 残差学习原理^[32]

残差学习的原理如图 3-8 所示。设网络的输入或上一层的输出为 x ，第一、二个变换层的权重分别为： W_1 、 W_2 ；记非线性变换 $\mathcal{F}(x) = \text{ReLU}(xW_1)W_2$ ，其中，ReLU 如式子(3-3)所示：

$$\text{ReLU}(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (3-3)$$

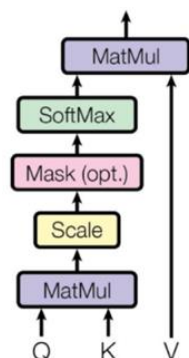
经过变换 $\mathcal{F}(x)$ 后，输出为 $\mathcal{H}(x)$ 。一般如 AlexNet、VGGNet、GoogLeNet 等网络直接学习网络的输出 $\mathcal{H}(x)$ 。但残差学习则在输入和输出层增加了一个输入 x 的恒等变换，使得网络的输出层 $\mathcal{H}(x) = \mathcal{F}(x) + x$ ，由于输入 x 无需学习，网络只需要学习 $\mathcal{F}(x)$ ，从而大大降低了学习的难度。 $\mathcal{F}(x)$ 也被称为残差如式(3-4)的所示：

$$\mathcal{F}(x) = \mathcal{H}(x) - x \quad (3-4)$$

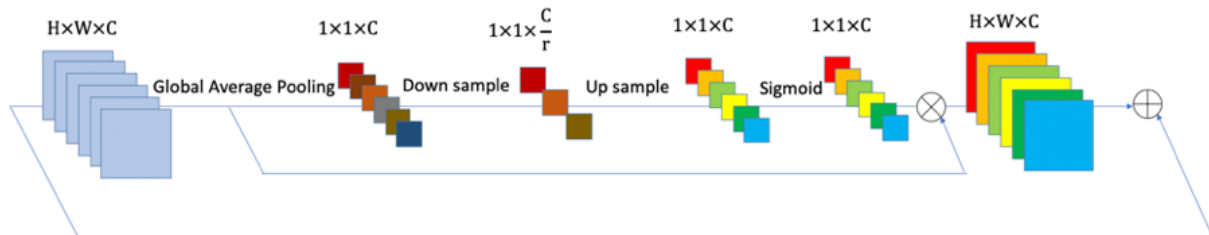
因此该学习原理被称为残差学习。由于残差学习在网络中加入了恒等变换使得网络在变深却没有性能提高的情况下，仍可通过恒等变换直连到下一层，即残差结果被学习为零，从而不影响网络性能。

2014 年，Google 学者 Ilya Sutskever^[33]等提出了 Seq2Seq 模型，该模型利用 LSTM 作为 Encoder-Decoder 模型的基本单元，完成输入的一种自然语言编码成固定长度大小并解码成为另一个种自然语言的任务。尽管该工作具有里程

碑式的意义，但也存在两个问题：第一，Encoder 的最后一个 hidden state 仅仅与句子末端信息联系紧密很难保证其能编码器句子起始部分；第二，Encoder 为每个句子赋予相同的权重，性能可能因此受到限制。为了解决以上问题，Dzmitry Bahdanau 等^[34]学者提出了注意力机制，希望对句子的词语实现重要程度区分。同年，Xu 等^[35]将注意力机制引入到计算机视觉，并提出了基于确定性得分计算的 Soft Attention 和基于概率采样计算得分的 Hard Attention。2015 年，Minh-Thang Luong 等^[36]学者改进了 Soft Attention 和 Hard Attention，提出了关注所有 Encoder 的 Global Attention 与关注局部 Encoder 的 Local Attention。通过以上叙述，注意力机制的原理可以总结如图 3-9 所示即：将 Query 与一组键值对(Key-Value)通过对 Query 和对应 Key 的适应度函数计算产生加权值并对 Value 加权和映射到输出。


 图 3-9 注意力机制原理图^[36]

ReCA 机制对图像的不同通道分配不同权重，实现不同通道的重要程度的区分。由此，该机制可作用在输入的不同等压平面及网络隐藏层中的不同通道，通过该机制以解决选择影响台风路径的高响应等压面的大尺度问题。ReCA 的原理如图 3-10 所示：首先，对输入特征图进行全局平均池化得到每个通道的描述；其次，对得到的通道描述进行倍率为 r 的下采样；再次，将上一次输出结果进行相同倍率的上采样，并利用非线性变换 Sigmoid 函数得到每个通道的注意力值；再者，对每个等压面与注意力值进行哈达玛乘积注意力值下的新等压平面；最后，整个过程通过残差学习的方式进行。


 图 3-10 具有残差特性的通道注意力机制原理图^[28]

ReCA 机制选择重要的等压平面并在不断学习的过程中强化其作用，且残差学习的方法使得网络的学习难度下降，即使在该层不同等压面没有区分度的情况下，也可以通过恒等变换与下一层直连不影响整个网络的性能。

3.3.3 台风三维时序表征的 Multi-ConvGRU 原理

LSTM提出后虽然解决了RNN模型无法学习长期依赖的问题，但LSTM复杂的门控结构却引入了过多参数、操作复杂和容易过拟合等问题。2014年，Chung等^[37]提出LSTM简化的Gated Recurrent Unit(GRU)，GRU将LSTM中的遗忘门与输入门合并成为了更新门，而输出门也被命名为了重置门。GRU的原理图如图3-11所示。

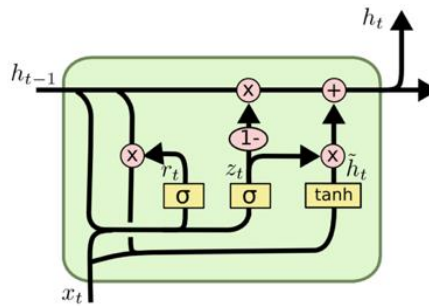


图3-11 GRU内部结构^[37]

在图中 z_t 表示更新门，更新门的作用是对上一时刻的输出和当前状态的信息进行控制， z_t 的值越大表示允许上一时刻的信息通过越多； r_t 表示重置门，用于对当前时刻到记忆单元 \tilde{h}_t 的信息控制， r_t 的值越小，表示允许传递的信息越少； h_t 表示GRU单元当前状态的输出。GRU单元的计算过程如下式(3-5)-(3-8)所示：

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (3-5)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (3-6)$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \quad (3-7)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (3-8)$$

上述(3-8)式展示了GRU输出的隐藏状态的计算过程。式中第一部分 $(1 - z_t) * h_{t-1}$ 通过加权比例控制上一时刻的信息流量，式中的第二部分 $z_t * \tilde{h}_t$ 则通过加权比例控制当前时刻的信息流量。在上一时刻与当前时刻的加权求和基础上实现GRU的输出。(3-8)式也实现了将LSTM的输入门和遗忘门进行合并，减少了参数、加快了训练速度并有效防止了参数过多而造成的过拟合问题。一般而言，在较少量数据的情况下，GRU的表现会好于LSTM。

在GRU的启发下，将ConvLSTM内部的LSTM单元换成了GRU单元，便得到了ConvGRU^[38]，该模型实现了时空序列预测的参数简化、加快训练速度。ConvGRU主要计算过程如式(3-9)-(3-12)所示：

$$Z_t = \sigma(W_{xz} * X_t + W_{hz} * H_{t-1} + b_z) \quad (3-9)$$

$$R_t = \sigma(W_{xr} * X_t + W_{hr} * H_{t-1} + b_r) \quad (3-10)$$

$$H'_t = f(W_{xh} * X_t + R_t \circ (W_{hh} * H_{t-1}) + b_h) \quad (3-11)$$

$$H_t = (1 - Z_t) \circ H'_t + Z_t \circ H_{t-1} \quad (3-12)$$

其中，*为卷积运算； \circ 为哈达玛乘积。

ConvGRU的提出虽然解决了时空序列学习的参数过多、学习速度慢等问题。但是，利用ConvGRU处理大气再分析数据，仍然存在特征提取不充分的问题。主要原因是传统的ConvGRU对输入图片和隐藏状态的进行一次卷积后执行GRU的操作。而大气再分析数据的精度不高，表示范围广泛，仅做一次卷积操作无法提取有效充分的特征。因此，本文提出了Multi-ConvGRU用于解决ConvGRU处理大气再分析数据的特征提取不足的问题。相比传统ConvGRU，Multi-ConvGRU则加入了多卷积模块，通过多次卷积实现非线性变换，达到提取更深层次的非线性特征的效果。多卷积模块对于精度相对不高的大气再分析数据而言，可以更好地提取其台风及其周边环境的特征，提高模型的准确率。Multi-ConvGRU的计算过程如公式(3-13)-(3-16)所示。

$$Z_t = \sigma(\gamma_{z_i}(X_t) + W_{hz} * H_{t-1} + b_z) \quad (3-13)$$

$$R_t = \sigma(\gamma_{r_i}(X_t) + W_{hr} * H_{t-1} + b_r) \quad (3-14)$$

$$H'_t = f(\gamma_{h_i}(X_t) + R_t \circ (W_{hh} * H_{t-1}) + b_h) \quad (3-15)$$

$$H_t = (1 - Z_t) \circ H'_t + Z_t \circ H_{t-1} \quad (3-16)$$

其中，*为卷积运算； \circ 哈达玛乘积， $\gamma_i(X_t) = W_1 * W_2 * \dots * W_i * X_t$ 为多卷积模块，本文中i被设定为3。

3.4 基于 wide&deep 异构模态台风路径特征融合模型

本小节将介绍台风二维路径时序特征与台风及其周边三维时序特征的异构模态融合方法。2016年，Google学者HT Cheng等^[39]提出了wide&deep特征融合的深度框架。论文指出，在数据相对稀疏的情况下，通过特征工程构建的广义线性模型与深度学习联合训练，将会比两个模型单独训练效果更好；论文认为通过特征工程构建的广义线性模型具有“记忆性”，可以记住某些特殊

关联，而由深度学习自动构建的特征具有“泛化性”，可以很好地推广，这两种特征融合在一起联合训练使得模型效果提高，图 3-12 是 wide&deep 融合模型的示意图。

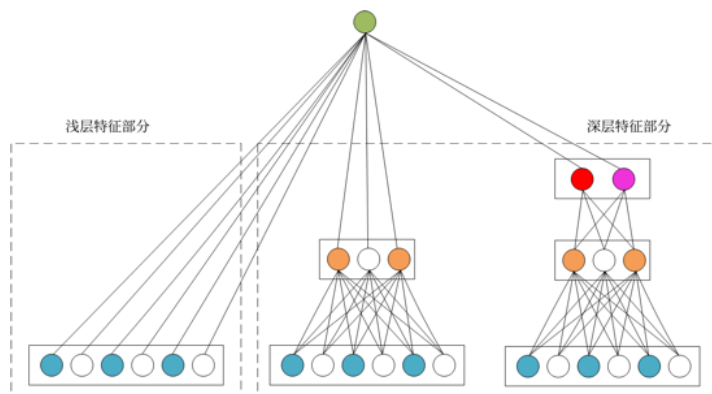


图3-12 wide&deep融合模型^[39]

基于 wide&deep 框架，将第 3.2 节提出的基于 CLIPER 的台风二维时序特征作为 wide 部分，该部分通过广义线性模型(Generalize Linear Model, GLM)建立一个具有记忆性的台风二维表征模型，并作为一个经验模型表征台风的已被认识的规律；GLM 是线性模型的扩展，通过联结函数建立响应变量的数学期望值与线性组合的预测变量之间的关系，可以很好地拟合台风的二维路径特征。其基本公式形式如下：

$$y = \mathbf{w}^T \mathbf{x} + b \quad (3-17)$$

其中， y 为预测值； $\mathbf{x} = [x_1, x_2, \dots, x_n]$ 是特征向量， n 为特征维度； $\mathbf{w} = [w_1, w_2, \dots, w_n]$ 是模型参数； b 为模型的偏置。

将第 3.3 节提出的基于 ReCA 机制和 Multi-ConvGRU 的台风及其周边三维时序特征作为 deep 部分，该部分主要通过深度学习建立一个具有泛化性的台风三维表征模型，可用于提取台风路径的深层未被认识特征。由于本章使用了两种台风的异构模态数据，因此两种特征融合之后，本章还为其增加了一个神经网络，作为两种特征的融合网络，具体数据流过程及网络结构如图 3-13 所示。

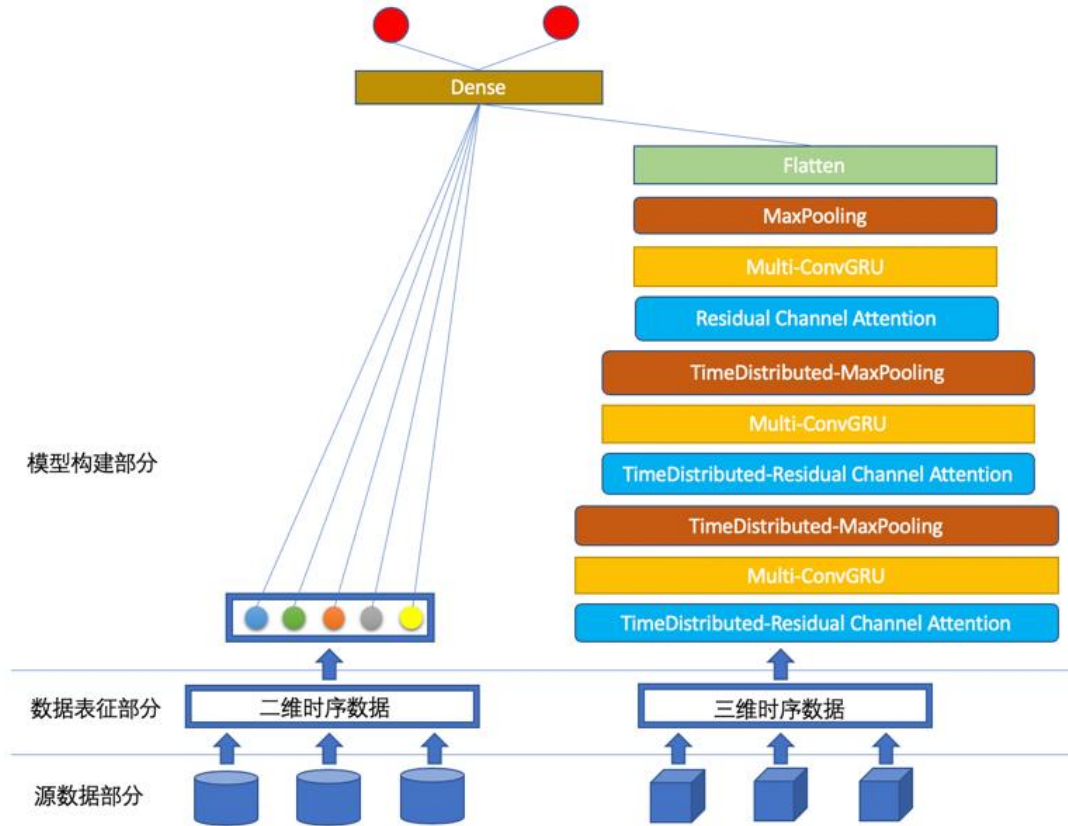


图3-13 数据流过程及网络结构

对于wide部分，根据CMA数据集构建了台风路径二维时序特征(见表3-2)，并采用GLM作为wide部分模型。对于deep部分，根据Geopotential变量，构建台风及其周边的三维时序结构。对构建好三维时序结构进行堆叠，以ReCA层、Multi-ConvGRU层、Max Pooling层三个网络层为一个堆叠块，共堆叠三次，其中前两次为多个时间维度上堆叠，最后一次对当前时间维度堆叠；网络最后将所有的特征做一次Flatten完成表征过程。得到wide部分与deep部分表征后，对这两部分以联合训练的方式后，增加一层神经网络融合两部分信息，最终得到台风24小时后经度与纬度的预测值。计算模型训练的损失，使用平均绝对误差(Mean Absolute Error, MAE)，计算预测的经度(\tilde{y}_{long})、纬度(\tilde{y}_{lat})与真实经度(y_{long})、纬度(y_{lat})之间的误差值：

$$\text{Loss} = \frac{1}{n} \sum_{i=0}^n (|y_{lat} - \tilde{y}_{lat}| + |y_{long} - \tilde{y}_{long}|) \quad (3-18)$$

3.5 实验结果与分析

3.5.1 实验环境

本章的台风路径预测模型实验环境如表 3-3 所示。

表 3-3 本章实验环境

名称	配置
操作系统	Ubuntu 16.04 LTS
CPU	Intel(R) Core (TM) i7-7800X CPU @ 3.50GHz
GPU	GeForce GTX 1080Ti*2
内存	64G
编程环境	Pytorch 0.4.1

3.5.2 实验设计

本章设计了三个实验以验证模型的性能。实验一，CMA 数据集 4 阶 CLIPER 特性分析实验，本实验目的在于验证 CMA 数据集是否具有 4 阶段 CLIPER 特性，是否适合使用 CLIPER 方法构建二维台风路径特征。实验二，GLM 有效性实验，本实验目的在于分析 GLM 模式是否比一般的机器学习、深度学习模型能够更有效表征台风的二维路径。实验三，基于 wide&deep 框架下异构模态融合有效性实验，本实验目的在于验证，基于 wide&deep 框架下，使用本文引入的 ReCA 机制和提出的 Multi-ConvGRU 方法是否对模型的性能有提高。

3.5.3 实验数据

本章使用了两种异源异构数据集包括：CMA 台风最佳路径数据集和 ECMWF 大气再分析数据集 ERA-interim 的 Geopotential 变量。CMA 数据集提供 1949 年以来西北太平洋(含南海，赤道以北，东经 180°以西)海域热带气旋每 6 小时的位置和强度，数据集内容包括：最佳路径记录时间、台风强度、中心经度、中心纬度、中心最低气压、2 分钟平均近中心最大风速。ERA-interim 数据集提供 1979 年以来全球大气每天 4 个时段(0 时、6 时、12 时、18 时)，囊括 37 个等压面，共 14 种大气再分析数据。由于早期台风记录数据集质量参差不齐，因此本章实验训练集选择为 2000-2014 年，验证集为该部分的 10% 的数据，测试集选择为 2015-2018 年。

3.5.4 实验结果与分析

(1) CMA 数据集 4 阶 CLIPER 特性分析实验

本实验根据 3.2.1 小节提出的方法，计算 CMA 数据集的 4 阶 CLIPER 特性，以判断 CMA 数据集是否适用 CLIPER 方法。实验结果如表 3-4 所示。

表 3-4 台风数据集 4 阶 CLIPER 特性

	CMA
台风总数	480
台风记录总数	14993
平均 4 阶 CLIPER 值	0.61
4 阶 CLIPER 值标准差	0.19
无持续性台风条数	9
无持续性台风记录总数	140
无持续性台风条数占比	1.88%
无持续性记录数占比	0.93%

表 3-4 实验表明：在 CMA 数据集上，台风的 4 阶 CLIPER 平均值显示台风具有较高的 CLIPER 特性，且不具有 CLIPER 特性的台风占比较少，符合使用 CLIPER 方法的条件（本文认为，若台风 4 阶 CLIPER 值小于零的台风记录数占比超过总数的 5%，则认为该数据集无持续性）。

(2) GLM 有效性实验

为了测试 GLM 作为 wide 部分的模型是否最佳，分别使用树型机器学习常用模型与深度学习常用模型进行实验并对比预测效果。实验结果如表 2-2 所示，使用式(2-14)作为评价指标。

表 3-5 CMA 数据集 wide 部分 24 小时预测距离(km)误差

	2015 年	2016 年	2017 年	2018 年	AVG
XGBoost	186.16	220.96	241.57	244.25	223.24
Random Forest	189.88	237.26	256.89	257.59	235.41
GBDT	181.24	213.75	232.12	235.48	215.65
1D CNN	171.15	213.18	220.10	215.98	205.10
LSTM	188.45	226.55	239.16	226.20	220.09
GRU	190.30	219.91	235.99	221.41	216.90
Simple RNN	211.26	247.35	278.30	267.99	251.23
GLM	172.03	211.89	215.16	210.80	202.47

表 3-5 实验结果表明：GLM 效果最佳。其中，树型机器学习算法预测效果不如 GLM 模型，原因可能是二维台风的主要特性是线性特点；常用深度学习模型效果不佳的原因可能是高质量二维台风的数据不够，深度学习无法学习到很好的台风二维特征。

(3) 基于 wide&deep 框架下异构模态融合有效性实验

为了测试融合的大气再分析数据是否有助于提高台风路径预测的精度，分别采样不同等压面 and 不同时间步使用 2D CNN^[56]、3D CNN、ConvLSTM^[57]、ConvGRU、Multi-ConvGRU 和 Multi-ConvGRU+ReCA 作为 deep 部分的表征。实验中单个等压面选取了 500hPa；四个等压面选取了 1000hpa、750hPa、500hPa、250hPa；八个等压面的实验选取了 1000hPa-300hPa(间隔 100hPa)。十六个等压面 1000hPa-250hPa(间隔 50hPa)。单时间步选取当前时间；多时间步选取了当前时间、前 6 小时、前 12 小时、前 18 小时；表 3-6 到表 3-9 分别给出了实验结果。实验使用式(2-13)作为评价指标。

表 3-6 单个等压面预测 24 小时距离(km)误差

	2015 年	2016 年	2017 年	2018 年	AVG
GLM	172.03	211.89	215.16	210.80	202.47
单时间步+GLM+2DCNN	154.62	150.87	163.36	181.95	162.70
多时间步+GLM+2DCNN	154.71	141.56	154.42	158.94	152.41

表 3-6 的实验结果表明：在单个等压面的条件下，加入大气再分析数据后采用 2DCNN 提取台风三维特征并与台风二维特征融合后预测的准确性有一定的提高。而且，多时间步的效果比单时间步的效果更好，原因是结合不同历史时间的台风三维结构对未来的预测更为有利。

表 3-7 四个等压面预测 24 小时距离(km)误差

	2015 年	2016 年	2017 年	2018 年	AVG
GLM	172.03	211.89	215.16	210.80	202.47
单时间步+GLM+2DCNN	139.99	161.06	155.05	171.57	156.92
多时间步+GLM+2DCNN	145.34	147.68	153.49	154.04	150.14
单时间步+GLM+3DCNN	164.04	153.33	169.50	154.29	160.29
多时间步+GLM+3DCNN	145.37	145.85	162.40	164.00	154.40
多时间步+GLM+ConvLSTM	141.40	147.51	156.38	167.34	153.16
多时间步+GLM+ConvGRU	138.23	140.56	145.38	166.80	147.74
多时间步+GLM+Multi-ConvGRU	129.18	146.37	154.24	148.36	144.54
多时间步+GLM+Multi-ConvGRU+ReCA	125.22	145.47	141.04	152.91	141.16

表 3-7 的实验结果表明：在四个等压面的条件下，实验基本结论与一个等压面的情况相同。3DCNN 在多时间步下得到的结果比 2DCNN 结果稍好，说明考虑台风的三维结构确实对路径的准确性提高有帮助。采用时空学习的方法 ConvLSTM 的效果不如 ConvGRU，原因是 ConvLSTM 参数过多容易造成模型过拟合。Multi-ConvGRU 的结果比 ConvGRU 的结果更有效是因为本文增加的多卷积模块提取了更广和更深的空间非线性特征。在 Multi-ConvGRU 的基础上加入 ReCA 后，模型效果进一步提升的原因是注意力机制能更加准确地建模了台风不同等压面下构成三维结构。

表 3-8 八个等压面预测 24 小时距离(km)误差

	2015 年	2016 年	2017 年	2018 年	AVG
GLM	172.03	211.89	215.16	210.80	202.47
单时间步+GLM+2DCNN	142.83	153.34	161.37	169.44	156.74
多时间步+GLM+2DCNN	135.41	142.61	152.16	157.04	146.80
单时间步+GLM+3DCNN	167.34	150.21	169.51	172.98	165.01
多时间步+GLM+3DCNN	143.96	142.23	154.44	158.22	149.71
单时间步+GLM+ConvLSTM	146.75	152.98	152.33	175.81	156.97
多时间步+GLM+ConvGRU	147.49	154.48	152.96	175.60	157.63
多时间步+GLM+Multi-ConvGRU	147.95	151.79	149.59	166.99	154.08
多时间步+GLM+ Multi-ConvGRU+ReCA	134.35	138.50	157.24	155.82	146.48

表 3-8 的实验结果表明：在八个等压面的条件下，实验结论基本与四个等压面相同。ConvGRU 的结果与 ConvLSTM 的结果相比变差，原因有可能是随着等压面的增加，更多参数的 ConvLSTM 能够表征复杂的三维台风结构。

表 3-9 十六个等压面预测 24 小时距离(km)误差

	2015 年	2016 年	2017 年	2018 年	AVG
GLM	172.03	211.89	215.16	210.80	202.47
单时间步+GLM+2DCNN	146.44	152.62	148.61	168.22	153.97
多时间步+GLM+2DCNN	137.48	141.41	151.78	152.00	145.67
单时间步+GLM+3DCNN	146.14	157.60	158.22	175.65	159.40
多时间步+GLM+3DCNN	138.20	140.70	154.90	154.09	146.98
多时间步+GLM+ConvLSTM	145.92	151.28	158.00	174.48	157.42
多时间步+GLM+ConvGRU	141.55	150.87	159.09	179.92	157.86
多时间步+GLM+Multi-ConvGRU	151.83	160.64	167.03	164.35	160.96
多时间步+GLM+ Multi-ConvGRU+ReCA	150.84	142.40	151.13	158.11	150.62

表 3-9 的实验结果表明：在十六个等压面的条件下，加入大气再分析数据后仍然可以提高台风路径预测的准确性。但是由于等压面的增多导致模型数据量增大，模型的训练效率和预测效率都有很大幅度的下降，且与其他更少等压面的结果相比，选取十六个等压面的预测结果都有不同程度的下降，且本文方法不再最优。主要原因可能是更复杂的等压面组合的三维台风结构使用简单的 3DCNN 提取其三维特征比复杂的方法更加有效。

最后，本实验总结如下：加入大气再分析数据构建台风及其周边三维时序结构后，台风路径预测的准确率有了明显的提升；对大气再分析数据加入时序效果与只使用当前时间步的预测效果相比普遍有提升；与 ConvLSTM 相比 ConvGRU 结果相对较好；使用提出的 Multi-ConvGRU 后，效果相对 ConvGRU 有提升，该提升得益于多卷积模块带来的多次非线性变换，能够更好地提取大气再分析数据的非线性特征；在 Multi-ConvGRU 基础上再加入 ReCA 机制后，效果得到进一步提升，该提升主要得益于通道注意力机制选取了代表台风前进方向的等压平面进行学习，并在残差学习的机制获得了较好的学习反馈；实验最后表明，在选取四个等压面的前提下，模型可以得到最优的效果，主要原因可能是等压面越多，越需要更复杂的模型表达台风及其周边三维时序结构，且增加越多的等压面会越容易造成模型过拟合现象。

3.6 本章总结

本章提出了利用深度学习算法解决台风路径预测问题的方法。首先，本章引入 CLIPER 特性检测方法，检测 CMA 数据集记录的台风路径是否具有 CLIPER 特性。接着，根据台风的 CLIPER 特性，利用台风的路径数据构建台风路径的二维时序特征。然后，利用大气再分析数据，构建台风及其周边的三维时序特征，并引入了 ReCA 机制对大气再分析数据的等压面分配不同的注意力值并提出 Multi-ConvGRU 模型处理精度不高的大气再分析数据。最后，使用 wide&deep 框架融合台风路径的二维时序特征与台风及其周边三维时序特征两种异构模态。实验证明，引入的 ReCA 机制和提出的 Multi-ConvGRU 模型，在 wide&deep 框架下进行异构融合模态，使得预测性能得到了提升。

第 4 章 基于多模态时空表示学习的台风强度预测

4.1 引言

台风强度是指台风中心的持续最大风速的大小,是判断台风所经过区域破坏强弱的重要指标。台风强度的预测是救灾资金估计与灾害损失估计的重要参考,在我国的抢险救灾事业发挥着重要的作用。

目前台风强度的预测方法各有优劣,主要分为 CNN 处理卫星图片^[40]和采用台风强度二维数据建模^[44]两种方法,这两种方法都存在对台风强度三维表征缺失的问题。且影响台风强度的因素主要来自台风内部的多尺度因素,因而台风的强度可视作是多尺度因素作用下的一种风速的合速度。由于多尺度因素复杂,直接预测合速度一般比较困难。因此,本文采用速度分量融合预测强度的方法解决此问题。为此,本文提出了多模态时空表示学习方法,融合台风强度二维和三维特征。在台风三维强度特征上,本章提出了一个空间注意力机制融合网络(Spatial Attention Based Fusing Network, SAF-Net)在空间注意力机制与卷积融合机制作用下,预测台风最终强度的模型。多模态时序特征表示学习方法是基于 wide&deep 框架的模型,其中 wide 表征部分通过改进台风路径二维时序特征,得到台风强度二维特征;deep 表征部分由包含两个组件的 SAF-Net 构成:一是空间注意力机制用于提取同一等压面风速分量高响应区域特征,二是卷积融合机制用于对风速分量单元的融合。

本章首先介绍基于 CLIPER 的台风强度二维时序表征算法;接着介绍基于 SAF-Net 的台风强度三维时序表征算法;然后介绍基于 wide&deep 异构模态台风强度特征融合模型;最后实验分析将用本文模型与传统方法进行对比,并实验证明空间注意力机制与卷积融合单元对模型效果有提升作用。

4.2 基于 CLIPER 的台风强度二维时序表征算法

根据第三章的实验结果,可知 CMA 数据集具有 4 阶 CLIPER 特性。台风强度信息表与表 3-1 相同,且构建台风强度的二维时序表征算法的约定与 3.2.2 小节相同,参考 Jin 等^[46]的研究成果构建台风强度二维时序特征如下表 4-1。

表 4-1 台风强度二维时序特征表

序号	特征	备注
1-5	LAT 、 LAT_6 、 LAT_{12} 、 LAT_{18} 、 LAT_{24}	前 24h 纬度
6-10	$LONG$ 、 $LONG_6$ 、 $LONG_{12}$ 、 $LONG_{18}$ 、 $LONG_{24}$	前 24h 经度
11-15	WND 、 WND_6 、 WND_{12} 、 WND_{18} 、 WND_{24}	前 24h 风速
16-20	$PRES$ 、 $PRES_6$ 、 $PRES_{12}$ 、 $PRES_{18}$ 、 $PRES_{24}$	前 24h 气压
21-24	$LAT - LAT_6$ 、 $LAT_6 - LAT_{12}$ 、 $LAT_{12} - LAT_{18}$ 、 $LAT_{18} - LAT_{24}$	前 6h-24h 一阶纬度差
25-28	$LONG - LONG_6$ 、 $LONG_6 - LONG_{12}$ 、 $LONG_{12} - LONG_{18}$ 、 $LONG_{18} - LONG_{24}$	前 6h-24h 一阶经度差
29-32	$WND - WND_6$ 、 $WND_6 - WND_{12}$ 、 $WND_{12} - WND_{18}$ 、 $WND_{18} - WND_{24}$	前 6h-24h 一阶风速差
33-36	$PRES - PRES_6$ 、 $PRES_6 - PRES_{12}$ 、 $PRES_{12} - PRES_{18}$ 、 $PRES_{18} - PRES_{24}$	前 6h-24h 一阶气压差
37-41	$\cos(LAT)$ 、 $\cos(LAT_6)$ 、 $\cos(LAT_{12})$ 、 $\cos(LAT_{18})$ 、 $\cos(LAT_{24})$	前 24h 纬度的 \cos 值
42-45	$\cos(LAT) - \cos(LAT_6)$ 、 $\cos(LAT_6) - \cos(LAT_{12})$ 、 $\cos(LAT_{12}) - \cos(LAT_{18})$ 、 $\cos(LAT_{18}) - \cos(LAT_{24})$	前 24h 纬度的 \cos 值的一阶差
46-50	$(WND)^2$ 、 $(WND_6)^2$ 、 $(WND_{12})^2$ 、 $(WND_{18})^2$ 、 $(WND_{24})^2$	前 24h 风速的平方
51-55	$\log(WND)$ 、 $\log(WND_6)$ 、 $\log(WND_{12})$ 、 $\log(WND_{18})$ 、 $\log(WND_{24})$	前 24h 风速的对数值
56-60	$(WND)^3$ 、 $(WND_6)^3$ 、 $(WND_{12})^3$ 、 $(WND_{18})^3$ 、 $(WND_{24})^3$	前 24h 风速的立方
61-65	$PRES/WND$ 、 $PRES_6/WND_6$ 、 $PRES_{12}/WND_{12}$ 、 $PRES_{18}/WND_{18}$ 、 $PRES_{24}/WND_{24}$	前 24h 气压与风速比
66-70	$WND/PRES$ 、 $WND_6/PRES_6$ 、 $WND_{12}/PRES_{12}$ 、 $WND_{18}/PRES_{18}$ 、 $WND_{24}/PRES_{24}$	前 24h 风速与气压比
71-75	$PRES/(WND)^2$ 、 $PRES_6/(WND_6)^2$ 、 $PRES_{12}/(WND_{12})^2$ 、 $PRES_{18}/(WND_{18})^2$ 、 $PRES_{24}/(WND_{24})^2$	前 24h 气压与风速平方比

表 4-1(续表) 台风强度二维时序特征表

序号	特征	备注
76-80	$(WND)^2/PRES$ 、 $(WND_6)^2/PRES_6$ 、	前 24h 风速与气压平方比
	$(WND_{12})^2/PRES_{12}$ 、 $(WND_{18})^2/PRES_{18}$ 、	
	$(WND_{24})^2/PRES_{24}$	
81-85	$PRES/(WND)^3$ 、 $PRES_6/(WND_6)^3$ 、	前 24h 气压与风速立方比
	$PRES_{12}/(WND_{12})^3$ 、 $PRES_{18}/(WND_{18})^3$ 、	
	$PRES_{24}/(WND_{24})^3$	
86-90	$(WND)^3/PRES$ 、 $(WND_6)^3/PRES_6$ 、	前 24h 风速与气压立方比
	$(WND_{12})^3/PRES_{12}$ 、 $(WND_{18})^3/PRES_{18}$ 、	
	$(WND_{24})^3/PRES_{24}$	
91-95	$\log(PRES)$ 、 $\log(PRES_6)$ 、 $\log(PRES_{12})$ 、 $\log(PRES_{18})$ 、 $\log(PRES_{24})$	前 24h 气压的对数值
96	Month	当前台风发生月份

表 4-1 中台风强度二维特征包括了台风二维结构的基本表征参数。其中，特征 1-36 为台风的线性信息，表示台风的线性结构：特征 1-20 是当前和历史台风基本信息，表征台风的当前和历史二维结构；特征 21-36 是当前和历史台风变化信息，表征台风的二维结构历史变化；特征 37-95 为台风的非线性信息，表征台风的非线性结构：特征 37-41 对台风的纬度做 \cos 非线性变换、特征 42-45 对纬度 \cos 非线性变换后求一阶差、特征 46-60 对风速做平方、对数、立方非线性变换；特征 61-95 为计算风速、风速的平方及风速的立方与气压的比值及该所有结果的倒数；特征 96 是台风发生月份，表征台风年度持续性的特点。

4.3 基于 SAF-Net 的台风强度三维时序表征算法

4.3.1 SAF-Net 中的空间注意力机制

台风强度预测问题与台风路径预测问题的三维结构主要区别是：第一，台风的路径是一种具有方向的矢量，而台风的强度不仅是一种具有方向的矢量，其更包含 U Wind 与 V Wind 两个速度分量的物理含义；第二，台风路径需要使用 Geopotential 描述的台风内部能量结构和地势高度，而台风的强度则需要使用 U wind 和 V wind 共同描述；第三，台风的路径是一个空间大尺度预测

问题，可以根据通道注意力机制进行学习其移动方向，而台风的强度是空间多尺度问题，可以根据空间注意力机制进行学习其力分布情况。空间域的注意力机制是将注意力集中在同一个特征图不同区域，以选择出每一个特征图的重要区域，为了更好地说明空间注意力机制，先介绍其发展历史。

由于传统 CNN 网络的池化层具有平移不变性，但其条件是特征平移小于池化矩阵的范围，可认为 CNN 只关注局部特性。2015 年，Google DeepMind^[47] 为克服该问题提出了 Spatial Transformer Networks(STN)，空间网络变换层。该网络结构具有平移不变性、旋转不变性及缩放不变性等特性，并且可以简易地加入到现有的卷积网络中，提高分类的准确性。该网络结构是属于空间注意力机制的一种，其目的是通过旋转等仿射变换操作，让网络结构更加注意空间中的高响应区域的学习。其原理图如图 4-1 所示。其中 Localisation net 用于生成仿射函数的参数，Grid generator 用于得到输出特征图的坐标点对应的输入特征图的坐标点的位置，Sampler 则是通过双向线性插值的方式得出输出特征图的某一点灰度值。

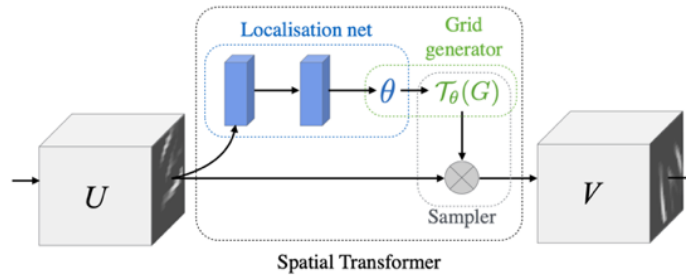


图 4-1 Spatial Transformer 原理图^[47]

2017 年，Wang^[48]等人利用空间注意力机制，提出 Non-local Neural Networks，该网络结构主要为了改进传统 CNN 感受野的局部问题。传统 CNN 网络的感受野范围可以通过卷积核得到，但是由于卷积核的大小有限，因此感受野的范围也会受限，虽然通过池化层后增加卷积会相应增大感受野，但池化层容易造成信息损失，从而导致感受野信息失真。为了解决这个问题，Non-local Neural Networks 引入了非局部的运算规则，如式(4-1)所示：

$$y_i = \frac{1}{C(x)} \sum_{\forall j} f(x_i, x_j) g(x_j) \quad (4-1)$$

其中， x 表示输入的信号， x_i 表示为该信号的某一个分量， f 是用于计算 x_i, x_j 相似性的函数， g 用于计算输入 x_j 在位置 j 上的特征值， $C(x)$ 为归一化参数。该网络提取的非局部特征流程如下：首先，先对输入的 X 做三次非线性变换得到压缩一半通道后的 θ, ϕ, g ；然后，计算 θ, ϕ 的相似性后得到注意力值；最后，

再与 g 相乘再重新扩大通道，得到空间注意力后的图像矩阵，整个学习过程通过残差学习的方式得到。网络的原理图如图 4-2 所示。

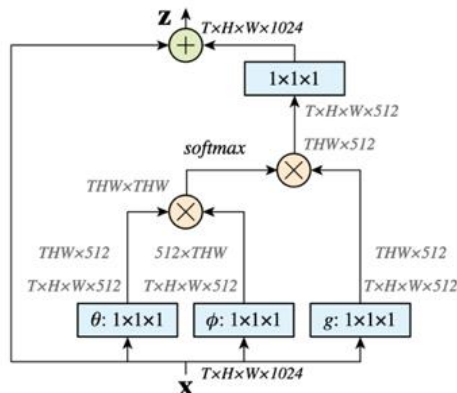


图 4-2 Non-local Neural Networks 原理图^[48]

2018 年，Sanghyun Woo^[49] 等人提出 Convolutional Block Attention Module(CBAM)，该工作建立在 Squeeze-and-Excitation Networks(SENNet)^[50]的基础上，SENNet 在特征图的通道上加入注意力机制，在不同的特征图中选择反馈能力强的特征图，而 CBAM 则保留 SENet 的通道注意力机制的基础上加入了空间注意力机制，更进一步提升网络模型的特征提取能力。该空间注意力机制原理如下：首先，网络会在不同的通道上进行平均池化；接着，对每一个通道最大池化后将这两个结果进行融合；然后，对特征图进行一次卷积核为 7×7 的卷积操作后，经过 Sigmoid 函数得到空间的注意力值；最后，将该空间的注意力值与原有特征图相乘，得到注意力机制后的新特征图。其原理如图 4-3 所示。

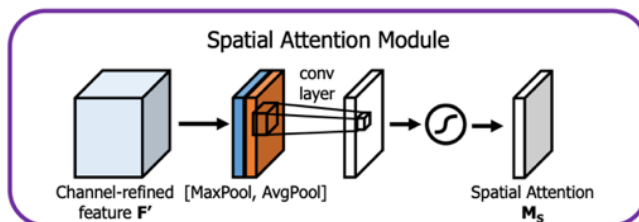


图 4-3 CBAM 中的空间注意力机制原理图^[49]

受到以上空间注意机制原理的启发，为学习不同等压面下 U wind 与 V wind 的注意力值，本文引入了另一种空间注意力机制^[52]。该注意力机制与 CBAM 相类似，其原理如下：首先，对输入的不同等压面做一次 1×1 的卷积操作，相当于对所有等压面做了一次非线性变换，得到结果后进行批归一化和 ReLU 非线性变换；然后，对结果再做一次 1×1 的卷积操作，得到结果后进行最后一次批归一化，通过 Sigmoid 函数得到注意力值；最后，利用得到的注意力值与原

等压面进行哈达玛乘积，得到经过注意力机制后的不同等压面的风速。本文引入的空间注意力机制原理图如图 4-4 所示。

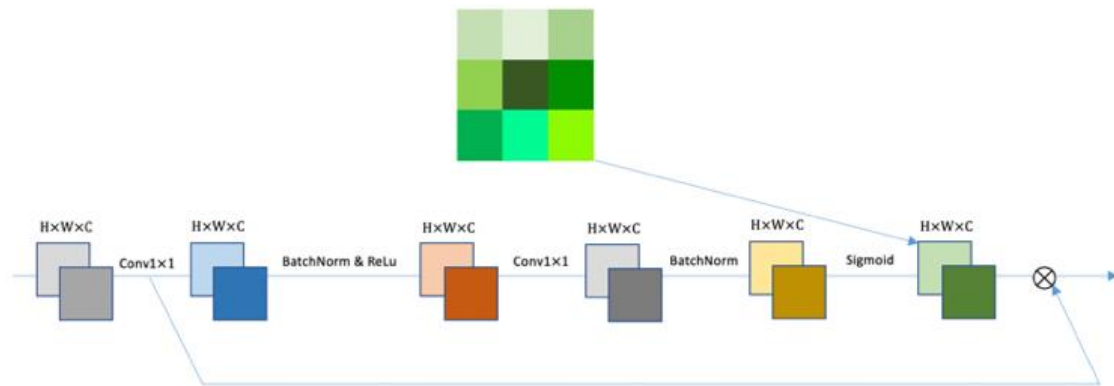


图 4-4 本文引入的空间注意力机制原理图^[52]

4.3.2 SAF-Net 中的融合机制

由于采用 U Wind 与 V Wind 两个分速度分别建模的方法，因此需要采用适当的融合策略。在机器学习和深度学习领域，一般通过特征融合的方法进行融合。

特征融合是指在模型的中间层，根据不同的特征特点进行数学组合，从而得到更好表达的特征的过程。根据刘渭滨等^[51]的研究，特征融合技术大致可以分为两种：基于机器学习的特征融合技术与基于深度学习的特征融合技术。

基于机器学习的特征融合技术主要包括：多个分类器的贝叶斯决策，该方法通过先验概率计算出后验概率，为不同的分类器计算融合权重，最后实现加性融合；利用稀疏表示矩阵进行特征融合，该方法通过对样本提取多种特征并建立特征联合稀疏表示矩阵，利用该矩阵完成多维特征选择，常见方法如 Lasso 回归。

基于深度学习的特征融合技术一般在隐藏层中实现特征融合，主要包括空间特征融合技术和时间特征融合技术。空间特征融合技术主要的融合对象是 CNN 的输出特征图，该方法通过对特征图进行一系列数学组合运算，完成特征的融合表达；时间特征融合技术主要的研究对象是 RNN 的输出序列结构或含时间信息的 CNN 输出的特征图，该方法通过对不同的时间序列输出进行池化，以完成特征融合。以下着重介绍深度学习的空间特征融合方法。

第一种融合方法是加性融合方法，该方法作用在两个具有相同大小、相同通道数的特征图上，利用数学的加法对两个特征图进行相加，其目的是为了突

出特征图中的高响应区域,但是由于线性相加会出现高响应区域与低响应区域中和的情况,可能会造成某些特征的信息损失,其融合公式如(4-2)所示:

$$Z = X^{h,w,c} + Y^{h,w,c} \quad (4-2)$$

其中, $X^{h,w,c}$ 和 $Y^{h,w,c}$ 分别表示两个特征图, h 表示特征图的高度, w 表示特征图的宽度, c 表示特征图的通道数。

第二种融合方法是最大融合方法,该方法作用在两个具有相同大小、相同通道数的特征图上,利用数学求最大值的方法对两个特征图求最大响应,其目的是为了突出特征图中的高响应区域,但是由于求最大值会使模型偏向高响应区域的表征,可能会造成对低响应区域的忽略,其融合公式如(4-3)所示:

$$Z_{i,j} = \max\{X_{i,j}^{h,w,c}, Y_{i,j}^{h,w,c}\} \quad (4-3)$$

其中, $X^{h,w,c}$ 和 $Y^{h,w,c}$ 分别表示两个特征图, h 表示特征图的高度, w 表示特征图的宽度, c 表示特征图的通道数; $X_{i,j}^{h,w,c}, Y_{i,j}^{h,w,c}$ 分别表示其对应特征图的第 i 行第 j 列的元素; 函数 \max 为求两个数最大值的函数。

第三种融合方法是通道融合方法,该方法作用在两个具有相同大小的特征图上但通道数可以不同,利用拼接的方法将两个特征图在通道的维度上实现合并,其目的是为了尽可能保留不同特征的响应区域,但是由于简单合并通道会导致通道数增加,可能无法得到很好的融合效果,其融合公式如(4-4)所示:

$$Z = \text{concatenate}\{X^{h,w,c_1}, Y^{h,w,c_2}\} \quad (4-4)$$

其中, X^{h,w,c_1} 和 Y^{h,w,c_2} 分别表示两个特征图, h 表示特征图的高度, w 表示特征图的宽度, c_1 表示特征图 X 的通道数, c_2 表示特征图 Y 的通道数; 函数 concatenate 为求将两个特征图在通道纬度上合并的函数。

第四种融合方法是卷积融合方法,该方法是:首先利用通道融合方法,得到通道合并结果后,然后对其进行卷积操作运算,其目的是为了利用卷积操作,对拼接的通道进行融合,其融合公式如(4-5)所示:

$$Z = f(\text{concatenate}\{X^{h,w,c_1}, Y^{h,w,c_2}\}) \quad (4-5)$$

其中, X^{h,w,c_1} 和 Y^{h,w,c_2} 分别表示两个特征图, h 表示特征图的高度, w 表示特征图的宽度, c_1 表示特征图 X 的通道数, c_2 表示特征图 Y 的通道数; 函数 concatenate 为求将两个特征图在通道纬度上合并的函数, 函数 f 为卷积操作函数。本文采用该融合方式对隐藏层的 U Wind 与 V Wind 进行融合。

第五种融合方法是外积融合方法,该方法作用在两个具有相同大小、相同通道数的特征图上,利用数学的外积方法对两个特征图(张量)求外积后再求和,

其目的是新的张量空间中求得与这两个特征图(张量)垂直的新张量，用于表达其两个特征图(张量)的融合结果，该方法会导致通道空间平方翻倍，急剧增加模型的复杂性。其融合公式如(4-6)所示：

$$Z = \sum_i \sum_j X_{i,j}^{h,w,c} \otimes Y_{i,j}^{h,w,c} \quad (4-6)$$

其中， $X^{h,w,c}$ 和 $Y^{h,w,c}$ 分别表示两个特征图， h 表示特征图的高度， w 表示特征图的宽度， c 表示特征图的通道数； $X_{i,j}^{h,w,c}, Y_{i,j}^{h,w,c}$ 分别表示其对应特征图的第 i 行第 j 列的元素； \otimes 表示外积运算。

4.4 基于 wide&deep 异构模态台风强度特征融合模型

本节将介绍基于 SAF-Net 的 wide&deep 异构模态台风强度特征融合模型。该网络结构的设计来源于 Liu 等人^[52]发表在 CVPR2019 的一篇论文，原论文网络结构利用空间注意力机制在共享全局特征池中提取不同任务所需要的特征，而本文则利用其空间注意力机制，提取不同特征池的特征并加以融合。本文提出的网络结构如图 4-5 所示。

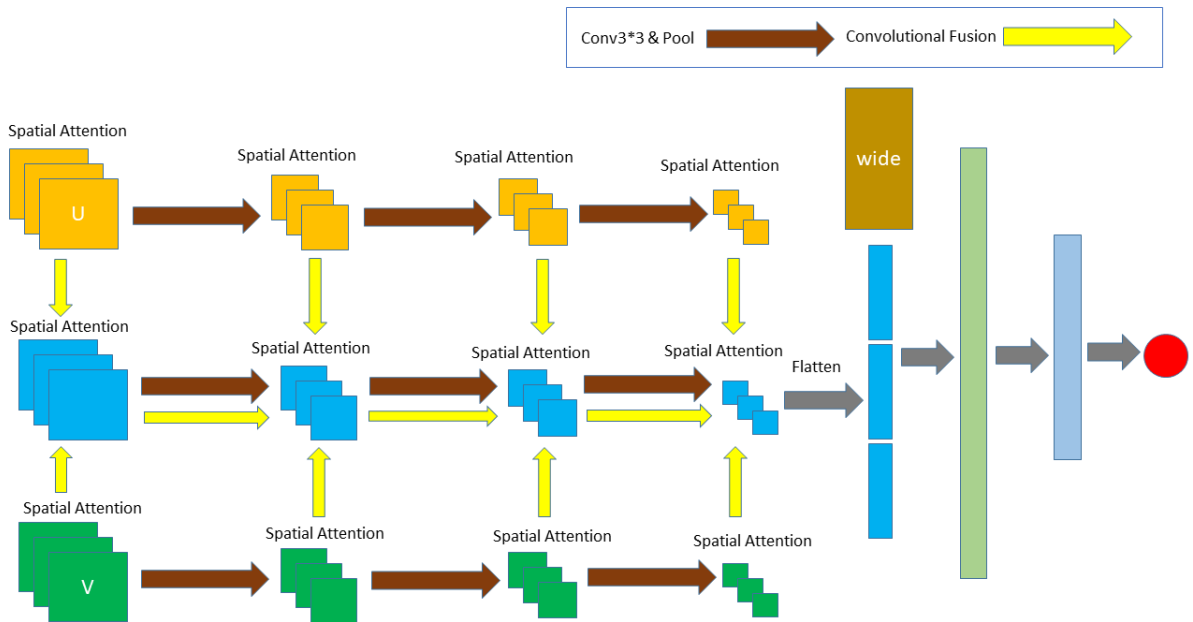


图 4-5 基于 SAF-Net 的 wide&deep 台风强度预测网络

对于wide部分，根据CMA数据集构建了台风强度二维时序特征(见表4-1)，并采用GLM作为wide部分模型。对于deep部分，根据U Wind与V Wind变量，构建台风及其周边的三维时序结构。本文提出的SAF-Net用于处理台风三维结构。如图4-5所示，SAF-Net的中间蓝色部分网络为主干网络，分别标示有U的

橙色部分和V的绿色部分的网络为枝干网络，且两个枝干网络结构相同，均用于提取台风分速度的特征。首先，分别对输入的U Wind和V Wind变量做一次空间注意力，得到两个输出后在主干网络进行卷积融合，得到特征图 R^1 ；接着，两个枝干网络同时做一次 3×3 的卷积操作和最大池化操作，所得结果进行空间注意力并在通道维度上进行融合后，得到特征图 R_2^2 ；然后，主干网络的特征图 R^1 与枝干网络进行相同的操作得到特征图 R_1^2 ，将 R_1^2 与 R_2^2 再进行一次卷积融合，得到主干网络的最新输出特征图 R^2 ；最后，枝干网络与主干网络重复第二步操作两次后，将最终结果做一次Flatten，完成deep部分的表征过程。得到wide部分与deep部分表征后，采用联合训练(jointly train)的方式，并增加一层神经网络融合两部分信息，最后得到台风24小时后强度的预测值。计算模型训练的损失，使用平均绝对误差(Mean Absolute Error, MAE)，计算预测的强度(\hat{y})与真实强度(y)之间的误差值：

$$\text{Loss} = \frac{1}{n} \sum_{i=0}^n |y - \hat{y}| \quad (4-7)$$

4.5 实验结果与分析

4.5.1 实验环境

本章的台风路径预测模型实验环境如表 4-2 所示。

表 4-2 本章实验环境

名称	配置
操作系统	Ubuntu 16.04 LTS
CPU	Intel(R) Core (TM) i7-7800X CPU @ 3.50GHz
GPU	GeForce GTX 1080Ti*2
内存	64G
编程环境	Pytorch 0.4.1

4.5.2 实验设计

本章设计了三个实验对模型进行验证。实验一，台风强度特征与台风路径特征对比实验，本实验目的在于验证 wide 部分台风强度二维时序特征与台风路径二维时序特征预测强度的有效性。实验二，SAF-Net 有效性实验，本实验目的在于验证 SAF-Net 的空间注意力机制与融合机制的有效性。实验三，基于

wide&deep 框架下异构模态融合有效性实验，本实验目的在于验证基于 wide&deep 框架下，使用本文提出的 SAF-Net 模型与现有方法对比的有效性。

4.5.3 实验数据

本章所使用的数据集 CMA 台风最佳路径数据集和 ECMWF 大气再分析数据集 ERA-interim 的 U Wind 变量与 V Wind 变量。训练集、验证集、测试集的划分方法与 3.5.3 节的方法相同。

4.5.4 实验结果与分析

(1) 台风强度二维时序特征有效性实验

本实验对表 3-2、表 4-1 分别采用 CLIPER 方法，预测 24 小时后的台风强度，判断表 4-1 作为预测台风强度的特征表的有效性。实验结果采用式(2-14)作为评价指标。

表 4-3 两种特征表 24 小时强度误差(m/s)

	2015 年	2016 年	2017 年	2018 年	AVG
表 2-2 的 CLIPER 方法	5.20	6.17	4.85	5.45	5.42
表 4-1 的 CLIPER 方法	5.15	6.44	4.77	4.98	5.34

表 4-3 实验结果表明：使用表 4-1CLIPER 方法构建的特征比直接使用台风路径表 2-2 CLIPER 方法更有效。该结果出现的原因可能是表 2-2 的 CLIPER 方法主要是将台风系统看作是一个大尺度问题，从台风整体出发建立特征。而表 4-1 的 CLIPER 方法则将台风系统看作是一个多尺度问题，从台风内部不同的风力状态出发构建特征。

(2) SAF-Net 有效性实验

本实验目的是测试 SAF-Net 中卷积融合机制和空间注意力机制的有效性。对于融合机制，将单独使用 U Wind 和 V Wind 的强度预测，并与两种变量融合进行对比，且融合机制将对比加性融合、最大融合与卷积融合；对于空间注意力机制，将在得到最佳融合方法后加入空间注意力对比模型是否有进一步提升。实验中单个等压面选取了 500hPa；四个等压面选取了 1000hpa、750hPa、500hPa、250hPa；八个等压面的实验选取了 1000hPa-300hPa(间隔 100hPa)。十六个等压面 1000hPa-250hPa(间隔 50hPa)。单时间步选取了当前时间；多时间步选取了当前时间、前 6 小时、前 12 小时、前 18 小时；实验结果如表 4-4 至表 4-8 所示。

表 4-4 单个等压面预测 24 小时强度(m/s)误差

	2015 年	2016 年	2017 年	2018 年	AVG
U+单时间步	5.48	6.40	4.45	5.21	5.39
U+多时间步	5.68	6.14	4.73	5.38	5.48
V+单时间步	5.32	6.02	4.95	5.26	5.39
V+多时间步	5.34	6.10	4.71	4.97	5.28
单时间步+加性融合	5.35	6.23	4.74	5.09	5.35
多时间步+加性融合	5.38	5.94	4.77	5.19	5.32
单时间步+加性融合+空间注意力	5.11	5.90	4.71	4.71	5.11
多时间步+加性融合+空间注意力	5.45	6.28	4.80	5.17	5.42
单时间步+卷积融合+空间注意力	5.12	5.85	4.61	4.80	5.10
多时间步+卷积融合+空间注意力	5.21	5.83	4.69	4.84	5.14
单时间步+最大融合+空间注意力	5.10	6.01	4.52	4.81	5.11
多时间步+最大融合+空间注意力	5.36	6.00	4.62	4.84	5.21

表 4-4 实验结果表明：在单个等压面的条件下，仅使用 U Wind 的预测效果比仅使用 V Wind 的预测效果要差，且采用加性融合效果也没有提升。单时间步的预测效果要好于多时间步的预测效果。出现这个现象的原因可能是所选择的单个等压面的影响。加入空间注意力后，实验结果有了较大的提升，说明空间注意机制可以很好地学习台风内部风速的重要区域。对于融合方法而言，加性融合效果最差、其次是最大融合，效果最好是卷积融合。

表 4-5 四个等压面预测 24 小时强度(m/s)误差

	2015 年	2016 年	2017 年	2018 年	AVG
U+单时间步	5.07	5.84	3.76	4.32	4.75
U+多时间步	5.01	5.38	4.30	4.43	4.78
V+单时间步	4.93	5.52	4.34	4.58	4.84
V+多时间步	5.06	6.12	4.51	4.61	5.07
单时间步+加性融合	4.72	5.27	4.01	4.13	4.53
多时间步+加性融合	4.67	5.35	3.99	4.10	4.53
单时间步+加性融合+空间注意力	4.52	5.15	4.00	4.13	4.45
多时间步+加性融合+空间注意力	4.43	5.17	3.80	4.03	4.36
单时间步+卷积融合+空间注意力	4.55	4.97	3.77	4.09	4.34
多时间步+卷积融合+空间注意力	4.54	4.78	3.95	3.94	4.30
单时间步+最大融合+空间注意力	4.55	4.87	3.89	4.01	4.33
多时间步+最大融合+空间注意力	4.63	5.09	3.89	4.13	4.44

表 4-5 实验结果表明：在四个等压面的条件下，融合后的效果比单独预测好。在加性融合基础上加入空间注意力后，效果得到进一步提高。说明在空间

注意机制作用下，融合后可以学习台风内部风速的重要区域。对于融合方法而言，加性融合效果最差、其次是最大融合，效果最好是卷积融合。

表 4-6 八个等压面预测 24 小时强度(m/s)误差

	2015 年	2016 年	2017 年	2018 年	AVG
U+单时间步	5.04	5.50	4.34	4.54	4.86
U+多时间步	5.18	5.48	4.20	4.55	4.85
V+单时间步	4.93	5.57	4.51	4.89	4.98
V+多时间步	5.13	5.55	4.24	4.84	4.94
单时间步+加性融合	4.49	5.01	3.80	4.44	4.44
多时间步+加性融合	4.79	5.14	3.85	4.43	4.55
单时间步+加性融合+空间注意力	4.59	5.08	3.84	4.22	4.43
多时间步+加性融合+空间注意力	4.57	5.12	4.06	4.05	4.45
单时间步+卷积融合+空间注意力	4.36	5.15	4.15	4.05	4.43
多时间步+卷积融合+空间注意力	4.41	4.98	4.01	4.05	4.36
单时间步+最大融合+空间注意力	4.70	4.82	3.99	4.07	4.39
多时间步+最大融合+空间注意力	4.67	5.00	4.13	3.93	4.43

表 4-6 实验结果表明：在八个等压面的条件下，实验结论基本与四个等压面的情况相同。但随着等压面增多，引入了不必要的噪声，导致结果较四个等压面的结果有不同程度的下降。

表 4-7 十六个等压面预测 24 小时强度(m/s)误差

	2015 年	2016 年	2017 年	2018 年	AVG
U+单时间步	4.68	5.34	4.26	4.72	4.75
U+多时间步	5.08	5.49	4.59	4.59	4.94
V+单时间步	4.96	5.71	4.53	4.84	5.01
V+多时间步	4.87	5.92	4.36	4.78	4.98
单时间步+加性融合	4.73	5.36	3.96	4.46	4.63
多时间步+加性融合	4.71	5.15	4.26	4.30	4.61
单时间步+加性融合+空间注意力	4.61	4.96	3.94	4.07	4.39
多时间步+加性融合+空间注意力	4.59	4.60	4.01	4.20	4.35
单时间步+卷积融合+空间注意力	4.58	4.94	4.04	4.07	4.41
多时间步+卷积融合+空间注意力	4.55	5.06	4.12	4.00	4.43
单时间步+最大融合+空间注意力	4.47	5.17	4.01	4.10	4.44
多时间步+最大融合+空间注意力	4.63	5.36	4.15	4.14	4.57

表 4-7 实验结果表明：在十六个等压面的条件下，融合后的效果依然比单独预测好。同样由于等压面的增多引入不必要的噪声，导致预测效果较八个等

压面的情况，也有着不同程度的下降。但空间注意力机制的作用还是较明显，加入后效果也有一定的提高。

最后，本实验总结如下：相比于只使用一种再分析数据(U Wind 或者 V Wind)采用融合两个分速度的方法效果最好；从融合的效果看来，当选取四个等压面构建台风及其周边三维时序结构得到的效果最好，所有实验(除十六个等压面外)，基本满足卷积融合效果最好，最大融合效果次之，加性融合最差的规律。加入空间注意力机制后，融合效果得到进一步的提高。

(3) 现有方法对比实验

在现有方法对比实验中，将使用本文提出方法与现有方法^[53]进行对比，对比分为两部分：主观预报方法对比与客观预报方法的区域模式对比。实验结果表明，本文提出方法平均效果优于目前现有方法。实验结果如表 4-8 所示。

表 4-8 现有方法 24 小时预测强度(m/s)误差对比

	2015 年	2016 年	2017 年	AVG
CMA	4.3	5.4	3.8	4.50
JMA	5.1	5.0	3.8	4.60
JTWC	4.9	5.0	5.0	4.97
KMA	5.1	5.7	4.0	4.93
澳大利亚数值	7.4	6.8	5.6	6.60
广州数值	8.9	8.0	6.5	7.80
上海台风模式	8.3	6.2	4.7	6.40
GRAPES-TCM	8.2	6.0	7.7	7.30
本文方法	4.54	4.78	3.95	4.42

4.6 本章总结

本章主要提出了使用深度学习算法解决台风强度预测问题。首先，根据台风的路径数据构建台风强度二维时序特征。然后，根据大气再分析数据构建台风及其周边三维时序特征，并引入了 SAF-Net 处理台风两种分速度。SAF-Net 主要由两部分组成：一部分为空间注意力模块，该模块可以通过空间注意力机制让模型更好地学习到台风的两个分速度的重要区域；另一部分为卷积融合模块，该模块通过将两个风速分量在通道维度上合并后，再通过卷积融合两部分风速。最后，使用 wide&deep 框架融合台风的二维与台风三维异构模态。实验证明，利用本章 SAF-Net 融合台风的 U Wind 与 V Wind 再分析数据，在 wide&deep 框架下，模型的性能得到了提升。

第 5 章 基于 Flask 的台风路径与强度可视化系统

5.1 引言

本章根据 CMA 数据集搭建了一个完整的台风路径与强度可视化系统，该系统展示了本文台风路径与强度的研究成果。本章首先将介绍该系统的整体设计框架及核心模块。然后详细说明系统各模块的功能及实现策略。最后介绍用户交互界面及展示系统运行效果。

5.2 系统架构设计

本文开发的基于 Flask 框架下的台风路径与强度可视化 Web 系统，从底层到顶层共分为：数据层、逻辑层和可视层。数据层提供数据支撑，存储 CMA 数据集与 predict CMA 结果集；逻辑层提供系统的各模块的具体实现；可视层面提供系统界面展示；完整的系统架构如图 5-1 所示。下面对系统架构图中包含的五个模块进行简要介绍：

数据存储模块：存储 CMA 数据集与模型预测结果的 predict CMA 结果集。数据采用逗号分隔值(Comma Separated Value， CSV)格式保存在服务器，数据库访问采用 Pandas 包。

历史台风路径与强度可视化模块：通过用户提供的台风信息，查询数据存储返回相应结果，并在地图上显示历史台风的路径与强度信息。

预测台风路径与强度可视化模块：根据用户所查询的历史台风，搜索预测台风信息，生成预测报告并可视化显示台风路径与强度。

模型对比模块：根据真实台风信息计算基线模型与本文提出模型的路径和强度误差，并可视化显示在 Web 系统上。

人机交互界面：提供给用户可视化操作界面，用户可以在页面选择相应条件，在地图上生成历史台风信息、预测台风信息。用户可查看各台风记录的具体信息和台风预测报告。系统前端使用 HTML5、jQuery、CSS3、Bootstrap 开发完成，后端使用 Flask、Python 开发完成。

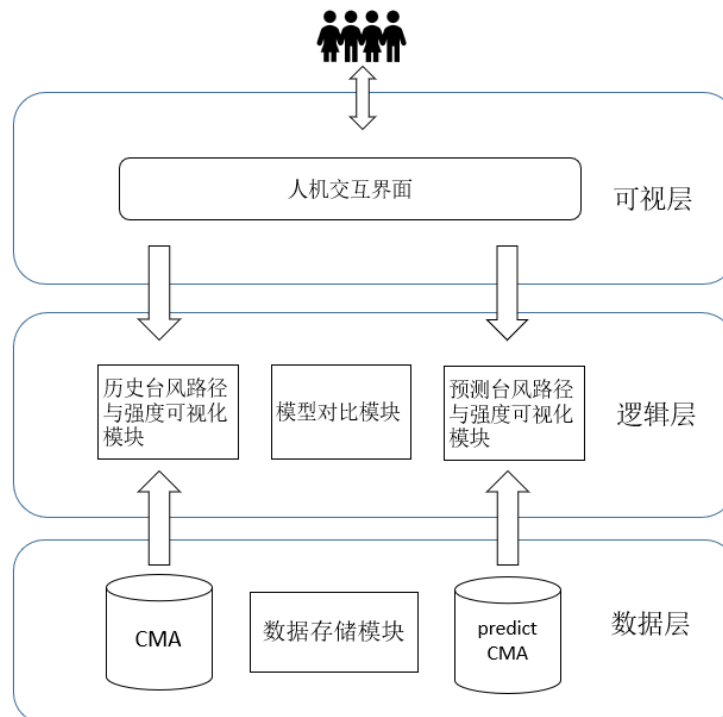


图 5-1 台风路径与强度可视化 Web 系统架构图

5.3 历史台风路径与强度可视化模块

历史台风路径与强度可视化模块需要用户在 Web 系统提供两个输入：台风发生的年份、所选年份台风名称。系统得到用户输入后，具体处理过程如下：可视层利用 Ajax 技术将用户输入传递到逻辑层；逻辑层根据输入，利用 Pandas 包连接数据层，访问 CMA 数据集并通过数据筛选函数取出符合用户选择条件的所有台风记录，所选择记录按时间顺序排列返回到逻辑层；逻辑层根据数据层返回的结果以 Ajax 方式回调输出；可视层利用 Ajax 请求成功后的回调函数获取逻辑层的返回结果，并将返回的字符串包装成用于显示的台风记录，其内容包括：台风发生时间、台风中心经度、台风中心纬度、台风中心风速、台风中心气压、台风强度。最后利用百度地图的 JS API 根据所提供的经纬度、并使用 jQuery 以动态的方式将台风路径显示在百度地图上，且用不同颜色的圆点标示台风强度的强弱。

5.4 预测台风路径与强度可视化模块

预测台风路径与强度可视化模块的输入与历史台风路径与强度可视化模块输入相同。系统获得的用户输入来源于历史台风路径与强度可视化模块，具

体数据处理流程如下：可视层利用 Ajax 技术将用户输入传递到逻辑层；逻辑层根据输入，利用 Pandas 包连接数据层，访问 predict CMA 数据并通过数据筛选函数取出符合用户选择条件的所有预测台风记录，所选记录按时间顺序排列返回到逻辑层；逻辑层根据数据层返回的结果以 Ajax 方式回调输出；可视层利用 Ajax 请求成功后的回调函数获取逻辑层的返回结果，并将返回的字符串包装成用于显示的预测台风记录，其内容包括：台风发生时间、台风中心经度、台风中心纬度、台风强度。最后利用百度地图的 JS API 根据所提供的经纬度、并使用 jQuery 以动态的方式将台风路径显示在百度地图上，且用不同颜色的圆点标示台风强度的强弱。

5.5 模型对比可视化模块

模型对比可视化模块的输入可分为两部分：历史台风路径与强度模块的输出与预测台风路径与强度模块的输出，即真实经度、真实纬度、真实强度与预测经度、预测纬度、预测强度。具体数据流程如下：用户选择台风发生年份、该年台风名后，分别点击查询和预测，可在可视层自动生成模型对比，展示内容包括：平均经度误差、平均纬度误差、平均距离误差、强度误差。

5.6 数据存储模块

数据存储模块主要存储 CMA 数据集与模型预测结果的 predict CMA 数据库。由于数据相对简单，因此 predict CMA 数据库仅采用 CSV 格式保存，且该数据库通过模型预测生成，具体表结构如表 5-1 所示。根据表 5-1 结构可知，本系统实现了模型预测过程与 Web 系统可视化过程的解藕。

表 5-1 predict CMA 数据表

字段	记号	备注
台风编号	TID	从 1 开始为每个台风编号，下一个台风自动增加 1
台风键值	KEY	组成为“台风发生年份-当年台风序号”
台风发生时间	TIME	日期格式为：YYYY-MM-DD-HH，每 6 小时记录一次
真实台风中心经度	LAT	北纬为正，单位是 0.1°N
真实台风中心纬度	LONG	东经为正，单位是 0.1°E
真实台风中心经度	PRED_LAT	北纬为正，单位是 0.1°N
真实台风中心纬度	PRED_LONG	东经为正，单位是 0.1°E
真实台风强度	WND	单位为 m/s
预测台风强度	PRED_WND	单位为 m/s

5.7 人机交互界面

将 Web 系统部署到服务器后，用户可以通过 IP+端口号的方式访问到台风路径强度可视化系统，本地部署可通过 127.0.0.1:8080 访问。网页首页如图 5-2 所示。网站首页提供台风发生年份下拉选择、台风名下拉选择，并有查询和预测按钮。页面左侧为基于百度地图的台风路径与强度展示地图，右侧为台风强度图例、历史台风路径信息、预测台风报告。

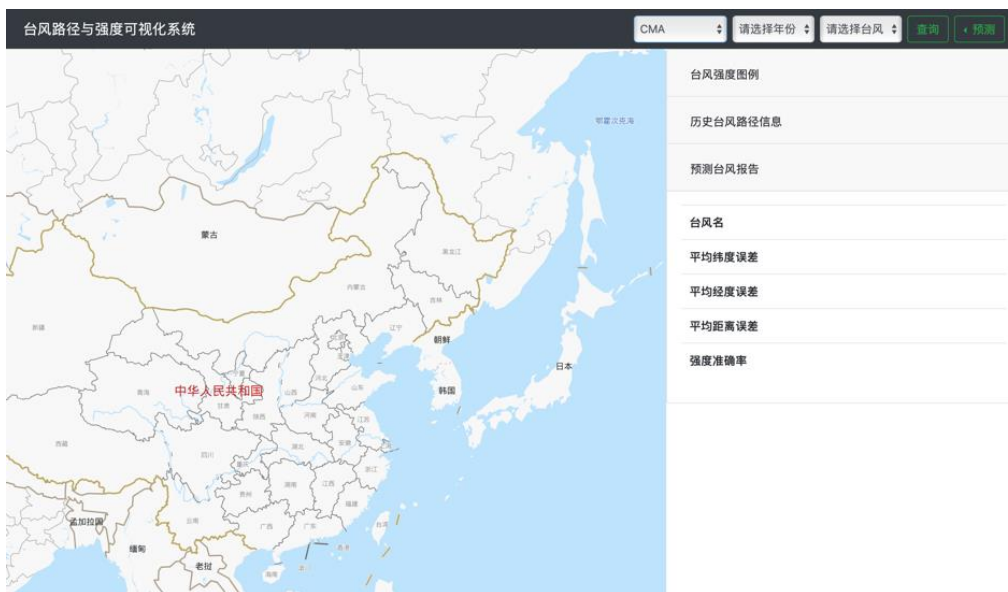


图 5-2 台风路径与可视化系统首页

用户点击“请选择年份”，可产生年份下拉列表，目前可选年份为 2000 年至 2018 年，如图 5-3 所示。

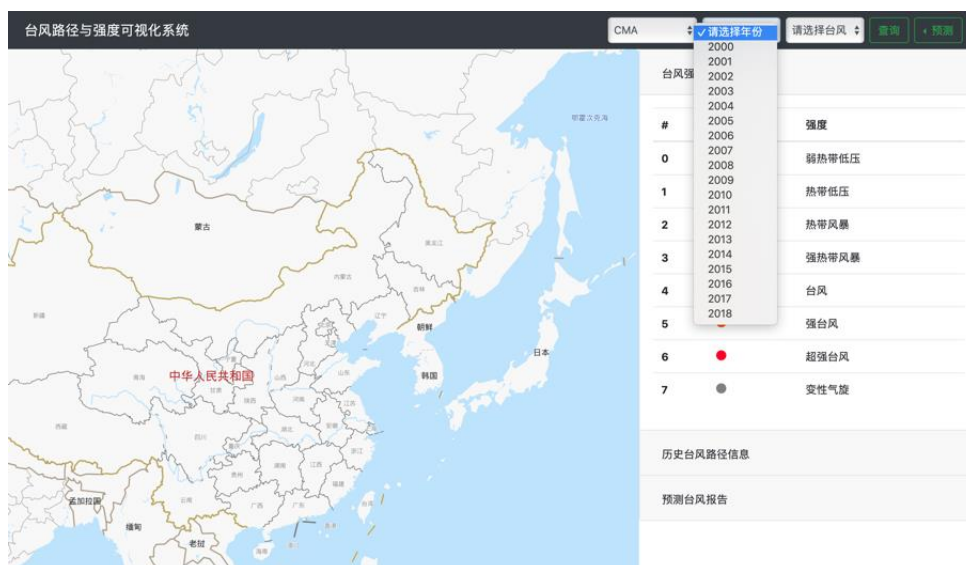


图 5-3 用户选择历史台风年份

用户点击“请选择台风”可得到所选年份的台风名，如图 5-4 所示。

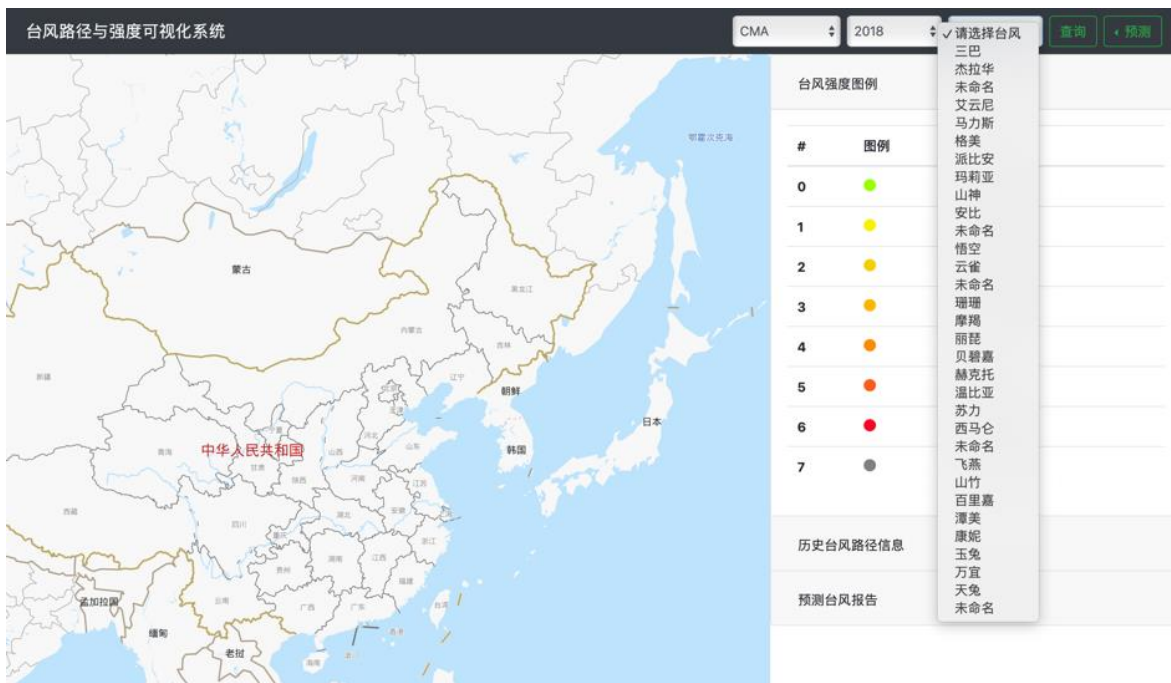


图 5-4 用户选择所属年份台风名

用户点击“查询”，Web 系统将动态生成台风路径并展示在左侧地图上，如图 5-5 所示。

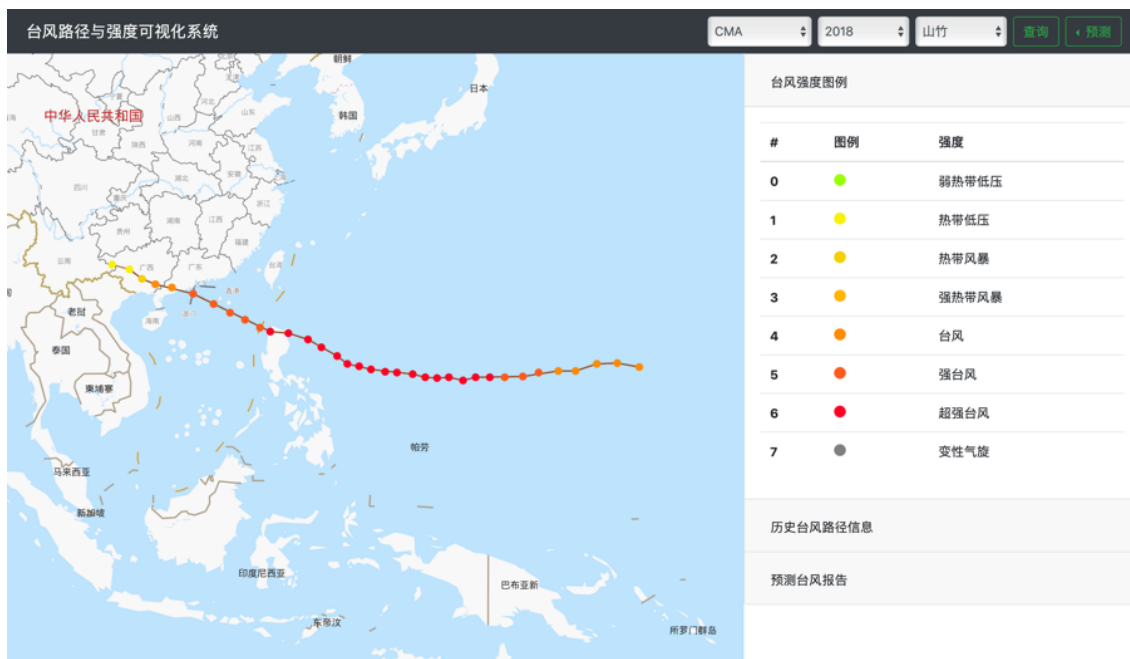


图 5-5 动态显示用户所选台风

用户点击右侧下拉栏目的“历史台风路径信息”，查看台风的历史信息。如图 5-6 所示。

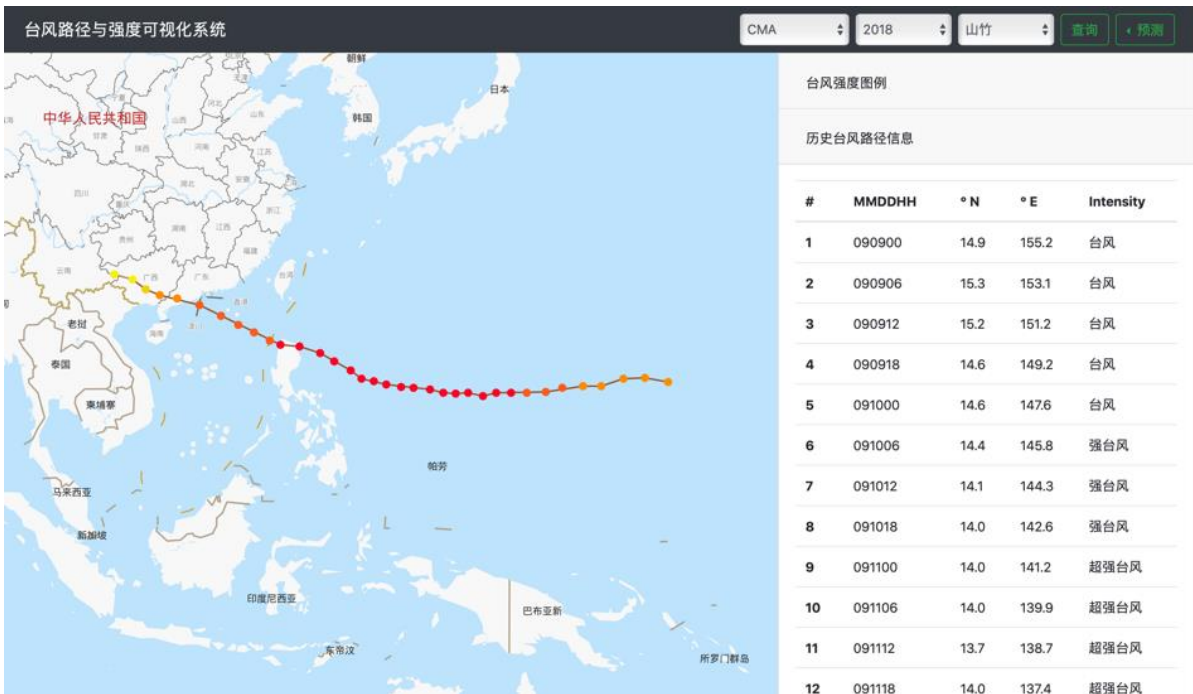


图 5-6 查看台风具体历史记录

用户点击右侧栏台风记录表的每一行或直接点击左侧地图的台风中心，可查看具体时间台风的具体信息。如图 5-7 所示。

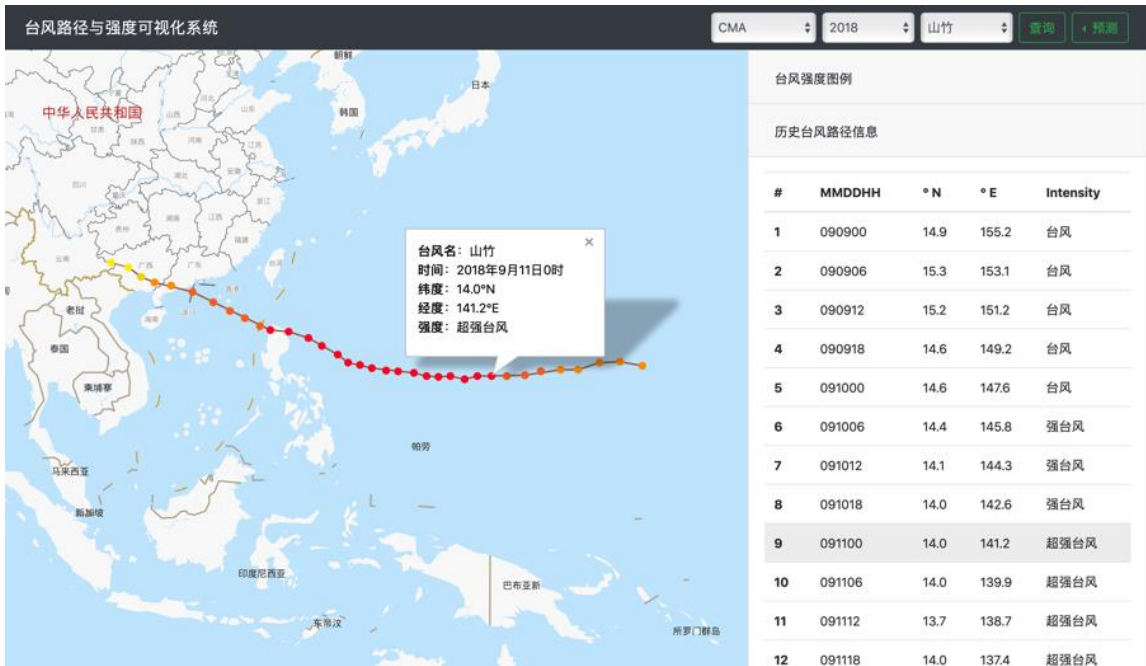


图 5-7 地图上查看台风具体历史记录

用户点击“预测”，选择“基线模型”，可以显示基线模型(灰色线)的预测结果。如图 5-8 所示。

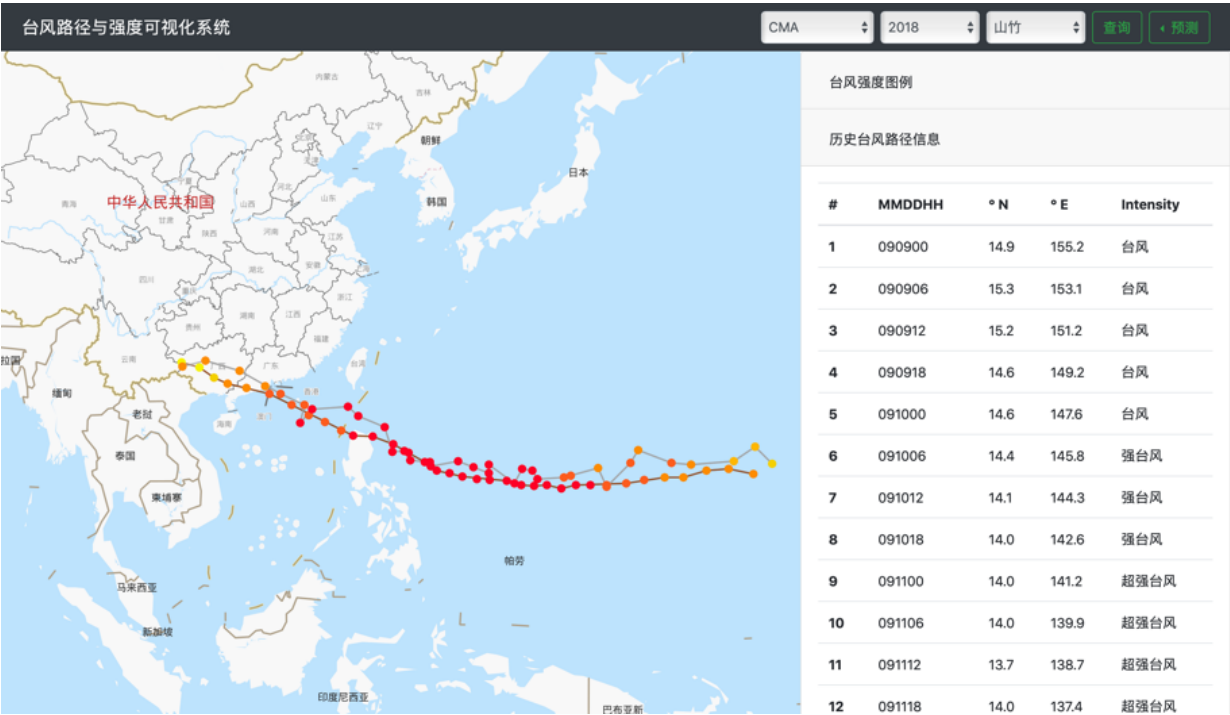


图 5-8 基线模型预测结果

用户点击“预测”，选择“我的模型”，得到本文提出模型(深蓝色线)的预测结果，点击“预测台风报告”可以对比本文提出模型(红色)与基线模型(黑色)对比结果。如图 5-9 所示。

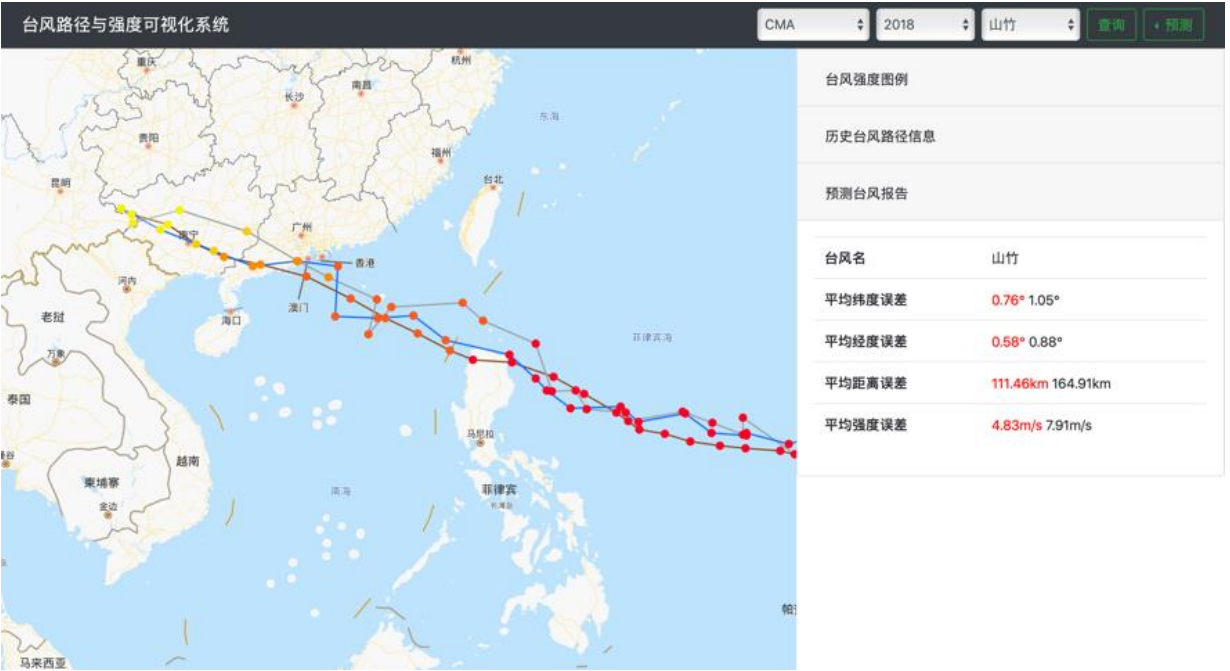


图 5-9 对比基线模型与本文提出模型预测结果

5.8 本章小结

本章主要介绍了基于 Flask 框架的台风路径与强度可视化系统的设计、功能实现和成果展示。从人机交互流程和系统开发的角度出发，介绍了本文设计的五个模块：数据存储模块、历史台风路径与强度可视化模块、预测台风路径与强度可视化模块、模型对比模块和人机交互界面。数据存储模块主要保存 CMA 数据集和模型生成的 predict CMA 数据库；历史台风路径与强度可视化模块主要作用是根据用户选择显示历史台风的路径与强度信息；预测台风路径与强度可视化模块主要作用是根据用户选择显示预测台风的路径与强度信息；模型对比模块根据实现基线模型与本文提出模型的对比报告与台风路径与强度的可视化展示；人机交互界面主要展示了本系统的可视化操作界面。整体来看，台风路径与强度可视化系统构建完成，但每个模块的细节还可以持续优化，未来可以进一步深入研究。

结 论

本文立足于台风路径与强度统计预测方法，在深度学习的基础上，研究了台风二维与三维的多模态时空序列表示学习问题，本课题的主要工作可以归纳为以下三个方面：

针对台风路径预测问题，针对模型影响台风路径等压面的选择问题，本文引入了 Residual Channel Attention 机制，充分考虑不同高度层的等压平面从而建立台风的三维立体结构。针对传统 ConvLSTM 无法提取大尺度非线性空间特征的问题，本文提出了 Multi-ConvGRU 为台风及其周边大气环境提取更广更深的非线性空间特征。最后实验结果表明：加入 ReCA 和 Multi-ConvGRU 后，台风路径预测效果较基线模型有明显提高。

针对台风强度预测问题，由于目前台风强度预测无法很好建模台风中心风场，本文将台风的中心风速利用 U Wind 和 V Wind 的大气再分析数据将其分解，并提出了 SAF-Net 结构学习融合两种分速度，该网络通过空间注意力模块关注每个等压平面的分速度高响应部分，通过融合模块将两个分速度融合成合速度。SAF-Net 采用了一个主干网络和两个枝干网络同时学习在线融合的办法，每一层网络学习完即刻融合，最后一层输出融合结果完成表征过程。实验结果显示：加入了 SAF-Net 网络实验效果有明显提高，对 SAF-Net 两个模块做消融实验后，效果下降。

针对台风路径与强度可视化问题，本文设计了一个基于 Flask 轻量级 Web 可视化系统，用于展示本文的研究成果。

虽然本文的研究有较好的结果，但是仍有不足之处和需要提高部分：

在台风路径预测任务上，虽然深度学习方法结合 CLIPER 方法有较大的提高，但相比较于目前实务的动力学方法仍有较大的差距。同时统计预测方法对以外的特殊情况考虑有限，可能未来需要引入台风行进规则强化预测效果。

在台风强度预测任务上，由于台风强度的迅速加强现象，导致台风强度的预测精度可能会迅速下降。这个现象发生原理复杂多样，模型无法完善地为此建模。同时，本文研究中也尝试结合台风强度等级和台风中心风速使用多任务方法进行研究，但是模型精度却因为多任务引入而下降，目前下降原因仍未清楚。

参考文献

- [1] Baltrušaitis T, Ahuja C, Morency L P. Multimodal machine learning: A survey and taxonomy[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 41(2): 423-443.
- [2] 易笑园, 余文韬, 闫智超, 等. 几种台风风暴潮预报方法在实际预报中的运用及比较[J]. 海洋预报, 2006, 23(4): 82-87.
- [3] Godske C L, Bjerknes V. Dynamic meteorology and weather forecasting[M]. American Meteorological Society, 1957: 13-18.
- [4] Reisner J, Rasmussen R M, Bruintjes R T. Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model[J]. Quarterly Journal of the Royal Meteorological Society, 1998, 124(548): 1071-1107.
- [5] 顾建峰, 殷鹤宝, 徐一鸣, 等. MM5 在上海区域气象中心数值预报中的改进和应用[J]. 应用气象学报, 2000, 11(2): 189-198.
- [6] 陈德辉, 薛纪善, 沈学顺. 中国气象局新一代数值天气预报系统 (GRAPES)[C]. 2006 年海峡两岸气象科学技术研讨会论文集, 2006, 39: 41-46.
- [7] Hong S Y, Park H, Cheong H B, et al. The global/regional integrated model system (GRIMs)[J]. Asia-Pacific Journal of Atmospheric Sciences, 2013, 49(2): 219-243.
- [8] 沈学顺, 苏勇, 胡江林, 等. GRAPES_GFS 全球中期预报系统的研发和业务化[J]. 应用气象学报, 2017, 28(1): 1-10.
- [9] 杜钧. 集合预报的现状和前景[J]. 应用气象学报, 2002, 13(1): 16-28.
- [10] Buizza R, Leutbecher M, Isaksen L. Potential use of an ensemble of analyses in the ECMWF ensemble prediction system[J]. Quarterly Journal of the Royal Meteorological Society, 2008, 134(637): 2051-2066.
- [11] Bishop C H, Toth Z. Ensemble transformation and adaptive observations[J]. Journal of the Atmospheric Sciences, 1999, 56(11): 1748-1765.
- [12] Neumann C J. An alternate to the HURRAN (hurricane analog) tropical cyclone forecast system[M]. Southern Region Headquarters, Scientific Services Division, National Weather Service. 1972, 1-24.

- [13] Aberson S D. Five-day tropical cyclone track forecasts in the North Atlantic basin[J]. Weather and Forecasting, 1998, 13(4): 1005-1015.
- [14] 廖木星, 黄立文, 胡基福. 在船舶航行条件下西北太平洋台风路径预报的 CLIPER 模式[J]. 中国航海, 1996, 2: 43-50.
- [15] Song H J, Huh S H, Kim J H, et al. Typhoon track prediction by a support vector machine using data reduction methods[C]// Proceedings of International Conference on Computational and Information Science. Springer, Berlin, Heidelberg, 2005: 503-511.
- [16] Wang Y, Zhang W, Fu W. Back Propagation (BP)-neural network for tropical cyclone track forecast[C]// Proceedings of 2011 19th International Conference on Geoinformatics. IEEE, 2011: 1-4.
- [17] Giffard-Roisin S, Yang M, Charpiat G, et al. Tropical cyclone track forecasting using fused deep learning from aligned reanalysis data[J]. Frontiers in Big Data, 2020, 3: 1-5.
- [18] Alemany S, Beltran J, Perez A, et al. Predicting hurricane trajectories using a recurrent neural network[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2019, 33: 468-475.
- [19] Kim S, Kim H, Lee J, et al. Deep-hurricane-tracker: Tracking and forecasting extreme climate events[C]// Proceedings of 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2019: 1761-1769.
- [20] Xingjian S H I, Chen Z, Wang H, et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting[C]// Proceedings of Advances in Neural Information Processing Systems. 2015: 802-810.
- [21] 洪梅, 张韧, 何金海. 基于动力统计模型重构的副热带高压中长期预报[J]. 气象学报, 2006, 64(6): 780-789.
- [22] 杨玉震, 王耀领, 胡邦辉, 等. 西北太平洋热带气旋强度统计动力预报的改进模型[J]. 海洋预报, 2010, 27(3): 1-6.
- [23] Zhang W, Vecchi G A, Villarini G, et al. Statistical–dynamical seasonal forecast of western North Pacific and East Asia landfalling tropical cyclones using the GFDL FLOR coupled climate model[J]. Journal of Climate, 2017, 30(6): 2209-2232.

- [24] Ying M, Zhang W, Yu H, et al. An overview of the China Meteorological Administration tropical cyclone database[J]. Journal of Atmospheric and Oceanic Technology, 2014, 31(2): 287-301.
- [25] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.
- [26] LeCun Y, Boser B, Denker J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1(4): 541-551.
- [27] Ji S, Xu W, Yang M, et al. 3D convolutional neural networks for human action recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 35(1): 221-231.
- [28] Zhang Y, Li K, Li K, et al. Image super-resolution using very deep residual channel attention networks[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 286-301.
- [29] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]// Proceedings of Advances in Neural Information Processing Systems. 2012: 1097-1105.
- [30] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [31] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 1-9.
- [32] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [33] Sutskever I, Vinyals O, Le Q V. Sequence to sequence learning with neural networks[C]// Proceedings of Advances in Neural Information Processing Systems. 2014: 3104-3112.
- [34] Cho K, Van Merriënboer B, Gulcehre C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]// Proceedings of Empirical Methods in Natural Language Processing. 2014: 1724-1734.

- [35] Xu K, Ba J, Kiros R, et al. Show, attend and tell: Neural image caption generation with visual attention[C]// Proceedings of International Conference on Machine Learning. 2015: 2048-2057.
- [36] Luong M T, Pham H, Manning C D. Effective approaches to attention-based neural machine translation[C]// Proceedings of Effective Approaches to Attention-based Neural Machine Translation. 2015: 1412--1421
- [37] Chung J, Gulcehre C, Cho K H, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. arXiv preprint arXiv:1412.3555, 2014.
- [38] Ballas N, Yao L, Pal C, et al. Delving deeper into convolutional networks for learning video representations[J]. arXiv preprint arXiv:1511.06432, 2015.
- [39] Cheng H T, Koc L, Harmsen J, et al. Wide&deep learning for recommender systems[C]//Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. 2016: 7-10.
- [40] Chen G, Chen Z, Zhou F, et al. A Semisupervised Deep Learning Framework for Tropical Cyclone Intensity Estimation[C]// Proceedings of 2019 10th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp). IEEE, 2019: 1-4.
- [41] Wang C, Xu Q, Zheng G, et al. Estimating Typhoon Intensity with Convolutional Neural Network[C]// Proceedings of IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2019: 10011-10014.
- [42] Kar C, Kumar A, Banerjee S. Tropical cyclone intensity detection by geometric features of cyclone images and multilayer perceptron[J]. SN Applied Sciences, 2019, 1(9): 2523-3971.
- [43] Chen B, Chen B F, Lin H T. Rotation-blended CNNs on a new open dataset for tropical cyclone image-to-intensity regression[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018: 90-99.
- [44] Yang S, Meng Q. Hurricane intensity prediction based on time series data mining[C]// Proceedings of 2019 10th International Workshop on the

- Analysis of Multitemporal Remote Sensing Images (MultiTemp). IEEE, 2019: 1-4.
- [45] Cloud K A, Reich B J, Rozoff C M, et al. A Feed Forward Neural Network Based on Model Output Statistics for Short-Term Hurricane Intensity Prediction[J]. Weather and Forecasting, 2019, 34(4): 985-997.
- [46] Jin Q, Fan X, Liu J, et al. Using eXtreme Gradient BOOSTing to Predict Changes in Tropical Cyclone Intensity over the Western North Pacific[J]. Atmosphere, 2019, 10(6): 341-361.
- [47] Jaderberg M, Simonyan K, Zisserman A. Spatial transformer networks[C]// Proceedings of Advances in Neural Information Processing Systems. 2015: 2017-2025.
- [48] Wang X, Girshick R, Gupta A, et al. Non-local neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7794-7803.
- [49] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3-19.
- [50] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7132-7141.
- [51] 刘渭滨, 邹智元, 邢薇薇. 模式分类中的特征融合方法[J]. 北京邮电大学学报, 2017, 40(4): 1-8.
- [52] Liu S, Johns E, Davison A J. End-to-end multi-task learning with attention[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019: 1871-1880.
- [53] 陈国民, 白莉娜, 万日金. 2015 年西北太平洋热带气旋预报精度评定[J]. 气象, 2017, 43(4): 501-507.
- [54] 陈国民, 张喜平, 白莉娜, 等. 2016 年西北太平洋和南海热带气旋预报精度评定[J]. 气象, 2018, 44(4): 582-589.
- [55] 陈国民, 张喜平, 白莉娜, 等. 2017 年西北太平洋和南海热带气旋预报精度评定[J]. 气象, 2019, 45(4): 577-586.

- [56] Liu Y, Racah E, Correa J, et al. Application of deep convolutional neural networks for detecting extreme weather in climate datasets[J]. arXiv preprint arXiv:1605.01156, 2016.
- [57] Kim S, Kim H, Lee J, et al. Deep-hurricane-tracker: Tracking and forecasting extreme climate events[C]// Proceedings of 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2019: 1761-1769.

哈尔滨工业大学学位论文原创性声明和使用权限

学位论文原创性声明

本人郑重声明：此处所提交的学位论文《基于深度学习的台风路径与强度预测方法研究》，是本人在导师指导下，在哈尔滨工业大学攻读学位期间独立进行研究工作所取得的成果，且学位论文中除已标注引用文献的部分外不包含他人完成或已发表的研究成果。对本学位论文的研究工作做出重要贡献的个人和集体，均已在文中以明确方式注明。

作者签名：徐光宇 日期：2020年7月3日

学位论文使用权限

学位论文是研究生在哈尔滨工业大学攻读学位期间完成的成果，知识产权归属哈尔滨工业大学。学位论文的使用权限如下：

(1)学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文，并向国家图书馆报送学位论文；(2)学校可以将学位论文部分或全部内容编入有关数据库进行检索和提供相应阅览服务；(3)研究生毕业后发表与此学位论文研究成果相关的学术论文和其他成果时，应征得导师同意，且第一署名单位为哈尔滨工业大学。

保密论文在保密期内遵守有关保密规定，解密后适用于此使用权限规定。本人知悉学位论文的使用权限，并将遵守有关规定。

作者签名：徐光宇 日期：2020年7月3日

导师签名：叶光明 日期：2020年7月3日

致 谢

时光飞逝，两年研究生生涯马上就要结束，谨此以毕业论文致谢向我人生路给予我帮助和鼓励的人们表示衷心的感谢。

非常感谢我的妻子和双方父母，在我压力最大的辞职考研期间依然给予我鼓励和支持，尤其是我的父母和妻子，他们的支持和理解让我在研究生入学考试的备考和毕业论文的研究过程中充满力量。

特别地，由衷地对我的恩师叶允明老师致以诚挚的感谢，感谢他以博大的胸怀接纳二本学历的我成为他的学生；感谢他在我申请推荐攻博的时候，给予我的信赖、支持与分享人生经验；感谢他在我毕业论文受挫，科研失败的时候对我的耐心倾听和体贴鼓励。感谢他在我的科研路上树立起一个负责任、有底线、有正气的研究者形象。

感谢憨厚可爱的李旭涛老师，在我毕业论文遇到问题时和叶允明老师一起指导我的研究，让我在他们的共同指导下努力攻克困难，完成研究，同时也感谢实验室的其他老师，教会我知识和给我人生建议。

感谢我本科同学，在我考研期间给予我精神上的鼓励和研究生就读阶段给我排忧解难，愿你们以后一切顺利安好。

感谢实验室的同学们，特别是姚远博士给我详细分享读博经验、做科研的经验；感谢罗楚耀博士在我的科研受阻时给予我无私的帮助；还有很多同学都在我的学习、生活给予我很大的帮助，在此我由衷表示感谢。

感谢我在深圳的姑婆和表姐们，对我生活无微不至的照顾和体贴，让我在异乡城市感受到家的温暖；特别感谢我的表叔，从我大学开始一直对我的人生和职业发展给予无私建议，并在我考研期间主动提出帮助。

最后，感谢疫情期间奋战在一线的医务工作者和抗疫工作者，是你们的无私奉献，才使得我可以安心在家完成毕业论文，愿此次疫情逝去的同胞们往生安息。