

# 基于Vision Transformer和多头注意机制的频率稳定性预测方法

杜东来, 荣娜, 张异浩

贵州大学, 电气工程学院, 贵州 贵阳

收稿日期: 2023年9月14日; 录用日期: 2023年11月9日; 发布日期: 2023年11月16日

## 摘要

电力系统频率稳定性预测(frequency stability prediction, FSP)对于快速准确地制定故障后控制措施具有重要意义。然而, 传统的数据驱动方法未能有效地将系统的时空特征纳入模型训练中, 并且存在系统信息利用不足、抗噪能力差和面对新拓扑的泛化能力差等问题。针对上述问题, 本文提出了一种基于多头注意机制和视觉变压器(Vision Transformer, ViT)的FSP预测方法。首先, 所提出的预测方法利用位置编码层来捕捉系统的拓扑结构, 以获得空间信息。其次, 通过使用多通道输入层来捕获系统时间序列数据以获得时序信息。然后, 采用基于多头注意力机制层来增强ViT模型的泛化能力和鲁棒性。在修改的新英格兰39总线系统上进行了测试。实验结果表明, 与传统深度学习方法相比, 本文所提ViT模型误差值在一般情况下降低了65.71%; 在30 dB、20 dB和10 dB噪声环境下降低了61.55%、56.61%和47.14%; 在拓扑改变环境下降低了49.09%, 这证明了所提模型具有更高的预测精度、更好的鲁棒性和泛化能力。

## 关键词

频率稳定性预测, 深度学习, 视觉变压器模型, 多头注意力机制

# Frequency Stability Prediction Method Based on Vision Transformer and Multi-Headed Self-Attention

Donglai Du, Na Rong, Yihao Zhang

Department of Electrical Engineering, Guizhou University, Guiyang Guizhou

Received: Sep. 14<sup>th</sup>, 2023; accepted: Nov. 9<sup>th</sup>, 2023; published: Nov. 16<sup>th</sup>, 2023

## Abstract

Frequency stability prediction (FSP) of power systems is of great significance for quickly and ac-

curately developing post fault control measures. However, traditional data-driven methods fail to effectively incorporate the spatiotemporal features of the system into model training, and there are problems such as insufficient utilization of system information, poor noise resistance, and poor generalization ability in the face of new topologies. In response to the above issues, this article proposes an FSP prediction method based on multi-head attention mechanism and Vision Transformer (ViT). Firstly, the proposed prediction method utilizes a position encoding layer to capture the topological structure of the system and obtain spatial information. Secondly, by using multi-channel input layers to capture system time series data to obtain temporal information. Then, a multi-head attention mechanism layer is used to enhance the generalization ability and robustness of the ViT model, tested on the modified New England 39 bus system. The experimental results show that compared with traditional deep learning methods, the error value of the ViT model proposed in this paper is generally reduced by 65.71%; Reduced by 61.55%, 56.61%, and 47.14% in 30 dB, 20 dB, and 10 dB noise environments; The reduction of 49.09% in the topology changing environment proves that the proposed model has higher prediction accuracy, better robustness, and generalization ability.

## Keywords

Frequency Stability Prediction, Deep Learning, Vision Transformer Model, Multi-Headed Self-Attention

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

电力系统频率特性揭示了发电与负载在功率失衡时的频率变化[1]。而随着电能消耗水平的不断提升以及大规模可再生能源的并网,电力系统的惯性逐渐降低,抵抗干扰的能力减弱[2]。在遭遇严重扰动时,甚至会可能出现频率崩溃现象,这将对社会活动及工业生产造成负面影响。因此,实时准确地进行在线FSP显得尤为关键。

传统的FSP方法可分为两类,即时域仿真和等效模型法。在[3]中,微电网被建模为 $d-q$ 参考系下的负反馈系统以预测孤岛多母线微电网的频率稳定性。在[4]中,建立了一个将调速器响应和频率偏差解耦的分析模型来预测最小频率。文献[5]建立了一个混合等效模型来研究频率响应性。混合模型包括等效转子模型、等效调速器模型和等效负载模型。文献[6]中构建了一个带有 $df/dt$ 惯性控制的简化3型风电模型以探索频率响应。尽管时域仿真逻辑严密,精度高,但其时间复杂度高。而等效模型简化组件模型以提高计算性能,但它降低了精度,因此,这两者都难以提供快速可靠的在线频率稳定评估[7]。

近年来,广域测量系统(wide area measurement systems, WAMS)和相量测量单元(phasor measurement units, PMU)的广泛应用,为基于机器学习的FS发展带来契机[8]。决策树[9]、支持向量机[10]、随机权重神经网络[11]被用于FSP。然而,此类浅层方法不足以应对电力系统的高维时变特性。因此,学者们引入了深度学习以弥补机器学习的短板,它在处理形如FSP的非线性问题上表现出了卓越的性能。长短时记忆网络(Long Short-Term Memory, LSTM) [12]和残差记忆网络[13]等方法侧重于提取时间信息,卷积神经网络(Convolutional Neural Network, CNN) [14]和图神经网络(Graph Convolutional Network, GCN) [15]等方法强势于融合空间特征。然而,仅考虑拓扑或时间信息来执行FSP并不够全面。因此,一些多模型组合[16] [17]被开发出来解决以上问题,例如, CNN-GRU 和 GNN-LSTM 等。但这些模型仍存在问题,例如,

在物理意义上同等对待时间与空间特征并不可靠，基于多模型的方法需要更长训练周期。

因此，本文提出了一种基于 ViT [18] 的 FSP 模型。所提出的 ViT 具有强大的时空特征提取能力。通过分块嵌入层、位置编码层、Transformer 编码器、全连接层将输入数据转换为其高级表示，以集成 FSP 的完整时空动态。因此，所提 ViT 的准确性、泛化能力和鲁棒性得到了提高。鉴于所提出的 ViT 的优点，本文利用它形成了一个深度学习框架，并将其应用于准确预测频率稳定性。此外，在噪声环境和系统拓扑变化情况下的实验也证明了该模型的优越性能。

## 2. 基于 Vision Transformer 的频率稳定评估框架

### 2.1. 样本生成

预测模型在修改后的新英格兰 39 节点系统上进行了测试，具体如图 1 所示。此系统涵盖了 13 台发电机，其中 9 台等效同步发电机和 3 台等效风力发电机。为了探究可再生能源并网对电力系统频率稳定性的影响，本文对原有的新英格兰 39 节点系统进行了一些调整。具体调整如下：新增了 3 个等效风电场，即等效风电场节点 40、41 和 42，将其分别与母线 2、29 和 39 相连。此外，为了探究不同工况对电力系统频率稳定性的影响，将部分可再生能源取代同步发电机，并调整等效风电场发电比例实现不同渗透率工况下的系统运行环境[19]。

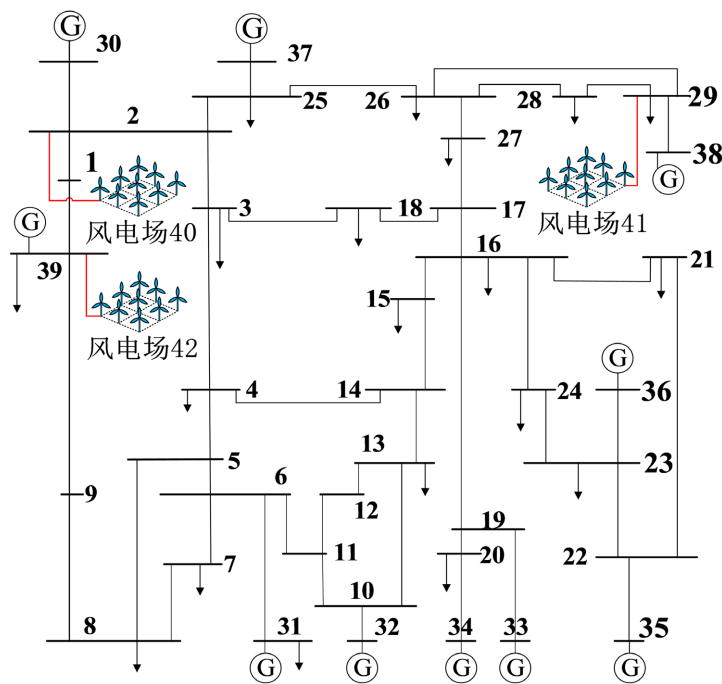


Figure 1. The modified New England 39-bus system wiring diagram  
图 1. 修改后的新英格兰 39 节点系统

实际电力生产中，电力数据的采集常受到各种干扰，因此测试所提出的模型的鲁棒性和适应性至关重要。本文进一步考虑了系统拓扑的改变以及系统受噪声影响的两种情况：1) 在基本实验数据集的故障轨迹后，分别引入了 10、20 和 30 dB 信噪比的白高斯噪声；2) 在 100% 负荷水平下引入有功扰动，并设置随机断开以生成拓扑变化样本，具体见表 1。上述关于测试系统的数值仿真实验由时域仿真软件 PSS/E 34 完成，本文将故障发生时间设置为仿真开始时间(0 秒处)，为故障发生后 60 秒处仿真结束，系统采样率为 100 Hz，间隔 2 秒在故障轨迹上采样故障点[20]。

**Table 1.** Description of samples generation  
**表 1.** 样本生成情况

基本实验	
负荷水平 $i$	80%, 90%, 100%, 110%, 120%
故障母线	30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42
有功扰动率	-40%, -38%, -36%, ..., 36%, 38%, 40%
渗透率	0%, 5%, 10%, ..., 30%, 35%, 40%
样本数	23,985
拓扑改变情况下的样本生成	
有功扰动率	-30%, -20%, -10%, ..., 10%, 20%, 30%
随机断线数	30
样本数	2730

## 2.2. 输入特征选择

在实际应用中, PMU、WAMS 潮流测量系统和电力系统时域仿真软件(如 PSS/E、DIgSILENT)被广泛使用, 这使得深度学习的模型输入特征可被获取。而频率稳定预测是一个典型的多输入时序回归任务, 并通过深度学习获取输入特征与频率稳定性的映射关系, 而原始特征的选择对于频率稳定性预测的准确性至关重要。合适的特征选择应当具有明确的物理意义, 输入特征维度越高, 系统状态被反映得越全面, 映射关系的准确度越高。但同时, 冗余的特征量会增大原始数据的获取难度并增加模型训练成本。因此, 本文参考了电力系统的突发干扰过程与过往研究进行输入特征的选择。

对于电力系统的突发干扰, 发电机基于相应的同步因子承受不平衡功率[21]。发电机节点  $j$  和节点  $k$  之间的同步因子表示如下:

$$SP_{jk} = V_j V_k (B_{jk} \cos \delta_{jk} - G_{jk} \sin \delta_{jk}) \quad (1)$$

其中  $V$  和  $\delta$  分别表示电压幅度和相角差,  $B_{jk}$  和  $G_{jk}$  表示传输阻抗。

而对于发电机的功率不平衡, 定义如下:

$$\Delta P = P_m - P_e = 2 \frac{H}{f_N} \frac{df}{dt} \quad (2)$$

其中  $P_m$  和  $P_e$  分别表示发电机的机械功率和电功率。  $H$  表示发电机的惯性常数。  $f_N$  表示系统运行频率。

而在电力系统的突发干扰过程中,  $V$ 、 $\delta$ 、 $P_m$  和  $P_e$  均是决定性的物理量, 综合考虑上述原因与相关文献[21] [22], 本文选择各母线、各发电机和各线路相关物理量的时间序列作为预测模型的输入特征, 具体输入特征如表 2 所示:

**Table 2.** Input data of models  
**表 2.** 型输入特征

序号	输入特征
1	每台发电机的有功功率
2	每台发电机的无功功率

## Continued

3	每个母线的有功功率
4	每个母线的无功功率
5	每个母线的电压幅值
6	每个母线的电压相角
7	每条线路的有功功率
8	每条线路的无功功率

### 3. 基于 Vision Transformer 的频率稳定预测方法

ViT 是一种在高维非线性预测领域中表现出卓越性能的深度学习模型。它基于 Transformer 架构，包括分块嵌入层、位置编码层、Transformer 编码器、全局汇聚层和全连接层。原始数据集被模型表示为一系列固定大小的块，每个块被展平成一个向量来作为模型的输入，再以全连接层输出所预测的系统频率轨迹。由于该架构可以提取输入特征中的位置信息和时序信息，契合于电力系统的频率稳定预测问题。

ViT 相较于传统深度学习方法具有一些优点，其通过 Transformer 编码器可以高效捕获全局特征，预防了深度学习的局部收敛问题。又通过多头注意力机制更好地利用输入数据中有价值的信息，并在一定程度上可以解释其预测过程，这使得模型决策过程变得更加透明。

#### 3.1. 分块嵌入与位置编码层

分块嵌入层将输入的原始数据  $x \in \mathbf{R}^{H \times W \times C}$  分成一系列固定大小的块  $x_p \in \mathbf{R}^{N \times (p^2 C)}$ ，每个块被展平成一个向量用作模型的输入，其过程类似于自然语言处理中的词嵌入，其中， $p^2$  是每个块的分辨率， $C$  代表输入数据的通道数， $(H, W)$  则代表输入数据的尺寸。而后，为了引入输入数据的位置信息，位置编码层被引入，由于 Transformer 没有显式的卷积操作，它无法捕捉到输入序列中元素的位置信息。因此，ViT 在嵌入向量中添加了位置编码，以便模型能够区分不同位置的特征，其公式通常如下：

$$\begin{cases} PE(pos, 2i) = \sin\left(\frac{pos}{10000^{2i/d}}\right) \\ PE(pos, 2i+1) = \cos\left(\frac{pos}{10000^{2i/d}}\right) \end{cases} \quad (3)$$

其中， $PE(pos, 2i)$  和  $PE(pos, 2i+1)$  分别表示位置  $pos$  上的奇数和偶数维度的位置编码， $i$  表示编码的维度， $d$  是嵌入的维度。

#### 3.2. Transformer 编码器

由位置编码层处理后的数据将由该编码器进行处理。而 Transformer 编码器是 ViT 模型的关键组件，其包含归一化层、多头注意力机制层和前馈神经网络层。得益于多头注意力机制模块，数据集集中的位置信息和特征间的关联信息在该机制的注意力权重下得以有效捕获。这有助于模型更好地理解图像中不同部分之间的关联性，从而提高模型的性能和泛化能力。多头注意力机制定义如下：

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

其中， $Q$  是查询(Query)矩阵， $K$  是键(Key)矩阵， $V$  是值(Value)矩阵， $d_k$  是键的维度。多头自注意力通过

多组不同的  $Q$ 、 $K$ 、 $V$  进行计算，然后将它们拼接在一起。

前馈神经网络通过多层的非线性变换，可以逐渐将经由多头注意力机制处理过后的输入数据映射到更高维的特征空间，从而生成信息量更大的特征表示，从而增加模型的适应性。因此，ViT 能对输入数据中更抽象、更高级别的特征进行建模，其定义如下：

$$FFN(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (5)$$

其中， $x$  是输入数据， $W_1$ 、 $b_1$  和  $W_2$ 、 $b_2$  分别是第一、二个全连接层的权重和偏置。

### 3.3. 全局汇聚与全连接层

ViT 中的全局汇聚是指数据在经过上述层处理后，在模型的最后阶段将特征向量从不同位置汇总成一个全局特征向量。这个全局特征向量可以用于后续的分类、回归或其他任务，定义如下所示：

$$Global\ Pooling(x) = \frac{1}{N} \sum_{i=1}^N x_i \quad (6)$$

其中， $N$  是特征的数量， $x_i$  是第  $i$  个位置的特征向量。

在这之后，全局汇聚后的特征向量被送入一个或多个全连接层，用于最终的回归任务。

## 4. 方法步骤

### 4.1. 数据标准化

为了降低输入特征维度和量纲对训练模型的影响，通过数据标准化对原始数据集进行预处理。这有助于使数据在不同特征之间具有更好的平衡，减少不同特征尺度之间的影响，从而提高模型的训练效果和收敛速度。此外，标准化也有助于防止梯度消失或爆炸等深度学习模型训练中的常见问题。

本文使用 Z-Score 归一化，也被称为标准化(Standardization)进行数据预处理[23]。此外，神经网络可能会忽略非归一化数据的分布特征，从而降低对其他数据的泛化能力。是一种常用的数据归一化方法，用于将数据转化为均值为 0，标准差为 1 的标准正态分布。这种归一化方法通过减去数据的均值，然后除以数据的标准差，将数据映射到一个具有零均值和单位方差的分布上，从而使得数据的分布更加稳定，标准化的计算方法如下所示：

$$X_{standardized} = \frac{X - \mu}{\sigma} \quad (7)$$

式中， $X_{standardized}$  是经过标准化后的数据点， $X$  是原始数据点， $\mu$  是数据均值， $\sigma$  是数据的标准差。

### 4.2. 基于 Vision Transformer 的频率稳定预测流程

基于 Vision Transformer 的频率稳定预测模型主要包括构建训练数据集、离线训练和在线应用 3 个部分，具体流程如下：

第一步：通过 PSS/E 34 建立所需新英格兰 39 节点系统模型；

第二步：对系统模型进行修改，新增三个等效风力发电机组，初始化模型参数，设置不同的切机状态与负荷水平  $i$ ，设置系统的运行方式；

第三步：进行数值仿真，记录深度学习的模型输入特征，构建训练数据集；

第四步：对模型进行超参数寻优，生成基于 ViT 的频率稳定预测模型；

第五步：将频率稳定预测模型离线训练好，并将模型保存，再进行线上应用。

具体过程如图 2 所示：

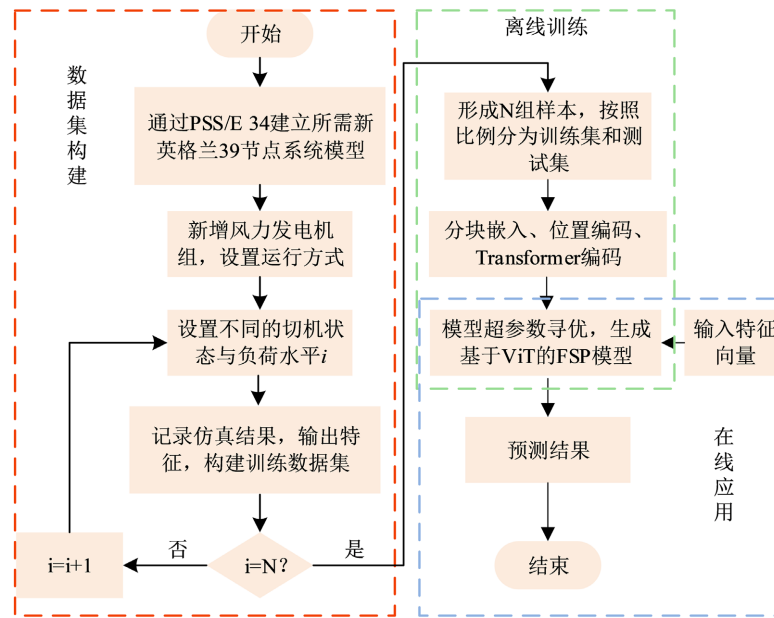


Figure 2. Flowchart of FSP based on VIT  
图 2. 基于 VIT 模型的频率稳定预测流程图

### 4.3. 模型评价指标

为了比较测试模型的性能, 本文采用平均绝对百分比误差(mean absolute percentage error, MAPE)和均方根误差(root-mean squared error, RMSE)来评价各模型的性能[24], 定义如下:

$$E_{MAPE}(f, I) = \frac{1}{m} \sum_{i=1}^m \frac{|f(x_i) - y_i|}{|f(x_i)|} \quad (8)$$

$$E_{RMSE}(f, I) = \sqrt{\frac{1}{m} \sum_{i=1}^m (f(x_i) - y_i)^2} \quad (9)$$

式中  $m$  是样本个数,  $i$  表示序号.  $x$  和  $y$  分别为样本的真实值与预测值.

模型损失函数采用 L2 范数损失函数. L2 范数损失函数, 即最小平方误差(LSE), 它将预测值与真实值的差值的平方和最小化, 其计算公式如下所示:

$$\text{loss}(x, y) = \frac{1}{n} \sum_{i=1}^n |y_i - f(x_i)| \quad (10)$$

式中  $n$  是样本个数,  $i$  表示序号.  $x$  和  $y$  分别为样本的真实值与预测值. 这些指标的数值越小对应着测试模型的性能越好.

### 4.4. 频率稳定指标

本文只考虑电力系统突然遭受有功功率扰动后的惯性响应和一次频率调节. 采用极值频率(extreme value frequency,  $f^{nadir}$ )、最大频率变化率(maximum rate of change of frequency,  $RoCoF^{max}$ )和准稳态频率(quasi-steady state frequency,  $f^{ss}$ )偏差来衡量频率稳定状态[22] [23] [24] [25]. 在扰动后的频率变化过程中, 各母线频率会围绕  $f_{coi}$  上下波动, 一般使用  $f_{coi}$  来表征系统频率, 其定义如下:

$$f_{coi} = \frac{\sum_{i=1}^n (H_i f_i)}{\sum_{i=1}^n H_i} \quad (11)$$

式中  $n$  表示电力系统中并网发电机组数量,  $H_i$  表征第  $i$  台发电机组的惯性时间常数,  $f_i$  代表第  $i$  台发电机组所在母线频率。

(1)  $f^{nadir}$ :  $f^{nadir}$  是  $f_{CoI}$  瞬时变化过程中的极值, 其决定切机或甩负荷的发生, 因而是频率扰动下最重要的指标。其值应满足  $f^{min} \leq f^{nadir} \leq f^{max}$ ,  $f^{min}, f^{max}$  是电网允许的最小/最大频率。

(2)  $RoCoF^{max}$ :  $RoCoF$  表征扰动后  $f_{CoI}$  变化快慢,  $RoCoF^{max}$  是故障后短时间内(如 500 ms)允许的最大频率变化率。为避免触发频率变化率保护装置,  $RoCoF^{max}$  应满足电网运行规范。

(3)  $f^{ss}$ :  $f^{ss}$  是系统中出现扰动后恢复到准稳态的频率值, 是检验系统恢复到准稳定状态的重要依据, 应满足电网运行规范。

## 5. 算例分析

### 5.1. 算例情况说明

仿真平台: Intel(R)Core(TM)i7-12700KF 3.61 GHz CPU、32 GB RAM 和 NVIDIA GeForce RTX 3070 GPU。所有涉及的算法都在 Pytorch-v1.10.1 和 Scikit-learn-v1.0.2 中实现。时域仿真软件为 PSS/E 34。

为验证所提方法的准确性、鲁棒性和适应性, 本文设置了三种情况对各模型进行测试: 1) 没有外界干扰情况下, 只考虑切机实验进行频率稳定性预测; 2) 考虑了系统拓扑频繁改变的情况, 以测试所提方法的适应性; 3) 考虑了原始系统数据受到噪声影响的情况, 以测试所提方法的鲁棒性。所测试模型包含 SVM、GRU、LSTM、ResNet、CNN、GCN 和 ViT, 所有模型均输入相同的训练集样本, 输入相同的测试集样本进行训练。在训练初始化时, 数据集被随机划分并采用五折交叉验证方法和 L2 损失函数验证性能, 测试模型均通过参数优化和多次训练达到了最小误差。参数优化之后, ViT 的学习率衰减率被设置为每 10 个 epoch 衰减到原来的 90%。激活函数和优化器分别设置为 GELU 和 Adam。初始学习率、权重衰减率和 epoch 分别设置为 0.01、0.0001、200。ViT 通过 Pytorch 框架实现, 其他模型通过 Scikit-learn 平台实现。ViT 的具体结构参数如表 3 所示:

**Table 3.** The parameters of ViT structure

**表 3.** ViT 结构参数

序号	超参数	值
1	输入维度	32
2	块尺寸	4
3	隐藏层维度	256
4	注意力头个数	8
5	前馈神经网络维度	128
6	Dropout 率	0.05

### 5.2. 各模型有效性比较

本节通过对性能指标的详细比较, 评估了各测试模型在频率稳定性预测方面的有效性, 相关数据可见于图 3。值得强调的是, ViT 模型在本节的比较中展现出了卓越的性能, 在测试集上, 其 RMSE 值为 0.0296, MAPE 值为 0.0289。这些误差指标在所有测试模型中均居于最低水平, 为 ViT 模型在 FSP 任务上的优越性提供了有力支持。相比之下, GCN 模型的性能较为逊色, 尽管 GCN 能够有效地提取系统的拓扑信息, 但其未能充分纳入时间信息进行模型学习。而 CNN 和 ResNet 模型的性能稍逊



于 ViT 和 GCN，在整体测试模型中处于中游水平，展现出不是很可靠的预测准确度，这是因为这两种模型侧重于捕捉系统的局部位置信息，然而对于电力系统的空间特征，并不能较为准确的表达。而 LSTM、GRU 和 SVM 模型的表现最不理想，因为这些模型并未融合拓扑信息，因此在 FSP 任务上的预测性能不尽如人意。从整体来看，ViT 模型的优越性得益于其多头注意力机制和全局特征提取的特点。这些特点允许模型充分挖掘频率稳定预测任务中的时空特征，使得其预测性能在实际应用中超越其他方法。

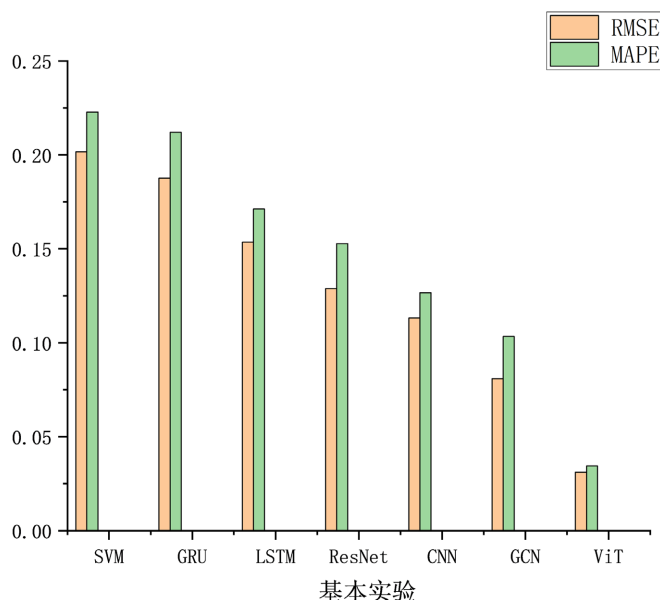


Figure 3. Basic experimental error chart  
图 3. 各模型性能对比

### 5.3. 噪声环境下各模型的鲁棒性比较

本节展示了各测试模型在噪声环境下的表现，具体指标数据见图 4。就测试系统而言，ViT 模型在 30 dB、20 dB 和 10 dB 噪声环境下的测试集的 RMSE 分别为 0.0311、0.0495 和 0.0684。从表中数据可以看出，随着噪声水平的增加，各模型的测试误差均呈现出上升的趋势。然而，在不同信噪比下，ViT 模型的预测精度始终优于其他模型。这一现象可以解释为，全局特征提取和多头注意力机制能够有效地过滤掉注意力与权重较低的特征，从而提高了模型的鲁棒性。与 ViT 类似，GCN 模型采用了切比雪夫滤波结构，但其未引入全局特征提取与多头注意力机制，因此其噪声过滤能力逊于 ViT。CNN 和 ResNet 模型同样具备滤波结构，但它们缺乏类似于 ViT 模型的多头注意力机制，因此它们的噪声过滤能力不及前两者。最后，LSTM、GRU 和 SVM 模型的性能表现最不理想，主要原因在于它们的浅层结构难以有效提取高维时空特征，同时也没有任何形式的滤波器来减弱噪声影响。因此，这些模型更容易受到噪声的干扰，进而影响了它们在噪声环境下的表现。综上所述，全局特征提取和多头注意力机制在 ViT 模型中的应用，使得该模型具备了较好的噪声过滤能力，尤其得益于多头注意力机制，模型的鲁棒性进一步提高，使得 ViT 模型在不同噪声环境下的表现均相对稳定。

### 5.4. 拓扑结构变化下各模型的适应性比较

本节展示了当测试系统的拓扑结构发生变化时，各模型的性能表现，具体数据见图 5。对测试系统而言，在拓扑变化发生时，尽管 ViT 模型的性能受到一定负面影响，但其在测试集上仍然保持着相对较

低的误差指标，其 RMSE 为 0.0511，MAPE 为 0.0584，这使其在测试模型中表现最佳。这种优越性归因于 ViT 模型的位置编码与多头注意力层。它们允许 ViT 模型将 FSP 的全面时空动态融入模型训练中，从而实现更高的性能。GCN 模型具有类似提取拓扑信息的图卷积结构，但由于未引入自注意力机制，其性能不如 ViT 模型。在拓扑结构变化的情况下，缺乏自注意力机制的 GCN 无法充分应对系统的变化，导致性能有所下降。与此同时，CNN 和 ResNet 模型的 RMSE 分别为 0.1162 和 0.1418。但值得注意的是，这两种模型所提取的矢量位置数据无法有效地反映拓扑结构的变化，因此它们的性能表现相对较差。而 LSTM、GRU 和 SVM 模型则没有任何提取拓扑信息的模块，导致它们在拓扑结构变化情况下的性能最为不佳。总的来看，图 5 的结果展示了 ViT 模型在处理系统拓扑结构变化时的相对优势。其有效的自注意力机制和位置编码设计使 ViT 模型能够更好地捕捉 FSP 的时空特征，从而在拓扑变化情况下保持较为稳定的性能。

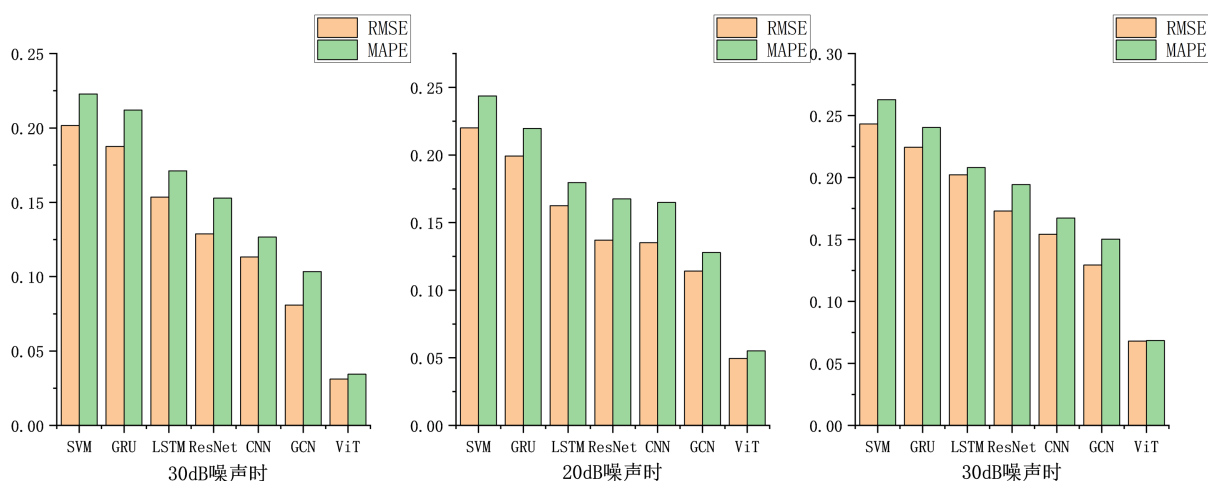


Figure 4. Performance comparison of various models under different noises

图 4. 不同噪声下各模型性能对比

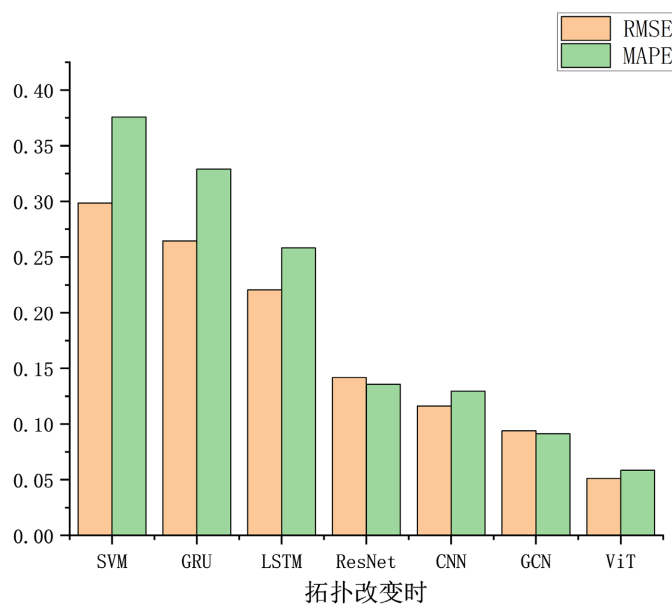


Figure 5. Comparison of performance among different models under topology changes

图 5. 拓扑改变下各模型性能对比

## 6. 结论

为了更好地将频率稳定预测的时空动态有机地融入学习模型, 本文提出了一种基于 ViT 的频率稳定预测方法, 克服了传统频率预测方法在准确性、鲁棒性和适应性方面的局限性。通过基于时域仿真软件对经过修改的新英格兰 39 节点系统进行了仿真研究分析, 并与其他深度学习算法进行了性能对比, 最终验证了所提方法的性能。根据实验结果得出以下结论:

1) 在 SVM、GRU、LSTM、ResNet、CNN 和 GCN 模型的测试中, GCN 的误差仅次于所提 ViT 模型并高于其余模型, 而本文所提 ViT 模型的误差值相较 GCN 降低了 65.71%。这证明在可再生能源, 即新增的 3 个风力发电厂并网后, 本文所提 ViT 模型在频率稳定性预测上具有更高的精度。

2) 采用位置编码和全局特征提取机制捕获了频率稳定预测的完整时空特征, 随后通过多头注意力机制持续关注所有输入特征的权重, 并通过 Transformer 联层聚合不同感受野的时空特征, 可以提高模型的抗噪能力和面对拓扑变换的适应性。实验结果表明, 对比所提出的 ViT 的误差值在 30 dB、20 dB 和 10 dB 噪声环境下降低了 61.55%、56.61% 和 47.14%; 在拓扑改变环境下降低了 49.09%, 这证明了所提模型在输入数据含有大量噪声时具有更好的鲁棒性, 在电力系统拓扑结构变化时具有更好的适应性。

## 基金项目

黔科合基础-ZK[2021]一般 277; 贵州省科技支撑计划项目([2023]290)。

## 参考文献

- [1] 林旻威, 温步瀛. 大规模风电接入对电力系统暂态稳定性影响研究综述[J]. 电气技术, 2017, 20(4): 1-8, 38.
- [2] 王琦, 郭钰锋, 万杰, 等. 适用于高风电渗透率电力系统的火电机组一次调频策略[J]. 中国电机工程学报, 2018, 38(4): 974-984.
- [3] 李锡林, 查晓明, 田震, 黄萌, 胡宇飞. 频率突变影响下基于 Lyapunov 法的孤岛微电网暂态稳定性分析[J]. 电工技术学报, 2023, 38(S1): 18-31+55. <https://doi.org/10.19595/j.cnki.1000-6753.tces.L10069>
- [4] 关中杰, 鲁效平, 李钢强, 等. 基于风速模型的风电机组动态转矩前馈控制技术[J]. 电工技术学报, 2018, 33(22): 5338-5345.
- [5] Azizi, S., Sun, M., Liu, G. and Terzija, V. (2020) Local Frequency-Based Estimation of the Rate of Change of Frequency of the Center of Inertia. *IEEE Transactions on Power Systems*, **35**, 4948-4951. <https://doi.org/10.1109/TPWRS.2020.3014818>
- [6] 陈宗遥, 卜旭辉, 郭金丽. 基于神经网络的数据驱动互联电力系统负荷频率控制[J]. 电工技术学报, 2022, 37(21): 5451-5461.
- [7] 苏玉刚, 颜志琼, 胡宏晟, 等. 基于频率切换实现恒流/恒压输出的电场耦合无线电能传输系统[J/OL]. 中国电机工程学报: 1-13. <https://doi.org/10.13334/j.0258-8013.pcsee.223014>, 2023-04-16.
- [8] Cao, Y., Zhang, H., Zhang, Y., Xie, Y. and Ma, C. (2020) Extending SFR Model to Incorporate the Influence of Thermal States on Primary Frequency Response. *IET Generation, Transmission & Distribution*, **14**, 4069-4078. <https://doi.org/10.1049/iet-gtd.2020.0206>
- [9] 石访, 张林林, 胡熊伟, 于之虹, 张恒旭. 基于多属性决策树的电网暂态稳定规则提取方法[J]. 电工技术学报, 2019, 34(11): 2364-2374.
- [10] 马良玉, 程善珍. 基于支持向量数据描述和 XGBoost 的风电机组异常工况预警研究[J]. 电工技术学报, 2022, 37(13): 3241-3249.
- [11] 叶瑞丽, 郭志忠, 刘瑞叶, 等. 基于小波包分解和改进 Elman 神经网络的风电场风速和风电功率预测[J]. 电工技术学报, 2017, 32(21): 103-111.
- [12] 崔昊, 冯双, 陈佳宁, 等. 基于自编码器与长短期记忆网络的宽频振荡广域定位方法[J]. 电力系统自动化, 2022, 46(12): 194-201.
- [13] 王彦博, 吴俊勇, 季佳伸, 等. 基于深度残差收缩网络的电力系统暂态频率安全集成评估[J]. 电网技术, 2023, 47(2): 482-492.

- 
- [14] 时纯, 刘君, 梁卓航, 等. 基于 GAN 和多通道 CNN 的电力系统暂态稳定评估[J]. 电网技术, 2022, 46(8): 3191-3202.
- [15] 王铮澄, 周艳真, 郭庆来, 等. 考虑电力系统拓扑变化的消息传递图神经网络暂态稳定评估[J]. 中国电机工程学报, 2021, 41(7): 2341-2349.
- [16] Xie, J. and Sun, W. (2021) A Transfer and Deep Learning-Based Method for Online Frequency Stability Assessment and Control. *IEEE Access*, **9**, 75712-75721. <https://doi.org/10.1109/ACCESS.2021.3082001>
- [17] Zhan, X., Han, S., Rong, N., *et al.* (2022) A Two-Stage Transient Stability Prediction Method Using Convolutional Residual Memory Network and Gated Recurrent Unit. *International Journal of Electrical Power & Energy Systems*, **138**, Article ID: 107973. <https://doi.org/10.1016/j.ijepes.2022.107973>
- [18] Dosovitskiy, A., Beyer, L., Kolesnikov, A., *et al.* (2020) An Image Is Worth  $16 \times 16$  Words: Transformers for Image Recognition at Scale. arXiv: 2010.11929.
- [19] 韩天森, 陈金富, 李银红, 等. 电力系统稳定评估机器学习可解释代理模型研究[J]. 中国电机工程学报, 2020, 40(13): 4122-4130.
- [20] 陈明华, 刘群英, 张家枢, 等. 基于 XGBoost 的电力系统暂态稳定预测方法[J]. 电网技术, 2020, 44(3): 1026-1033.
- [21] Yu, B., Yin, H. and Zhu, Z. (2017) Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. arXiv: 1709.04875. <https://doi.org/10.24963/ijcai.2018/505>
- [22] Zhao, J., Tang, Y. and Terzija, V. (2018) Robust Online Estimation of Power System Center of Inertia Frequency. *IEEE Transactions on Power Systems*, **34**, 821-825. <https://doi.org/10.1109/TPWRS.2018.2879782>
- [23] Lee, I. and Kang, J. (2019) Self-Attention Graph Pooling. *Proceedings of Machine Learning Research*, **97**, 3734-3743.
- [24] 赵荣臻, 文云峰, 叶希, 唐权, 李文沅, 陈云辉, 瞿小斌. 基于改进堆栈降噪自动编码器的预想事故频率指标评估方法研究[J]. 中国电机工程学报, 2019, 39(14): 4081-4092.
- [25] 文云峰, 赵荣臻, 肖友强, 刘赓斌. 基于多层极限学习机的电力系统频率安全评估方法[J]. 电力系统自动化, 2019, 43(1): 133-140.

## 附 录

表 A2~A5 为 PSS/E 中风力发电机的动态模型参数。表 A6~A8 为 PSS/E 中光伏发电机的动态模型参数。

**Table A1.** Parameters of WT3G2 module

**表 A1.** WT3G2 模块的参数

#1	#2	#3	#4	#5	#6	#7
0.20	0.0	0.0	0.0	0.10	1.50	0.50
#8	#9	#10	#11	#12	#13	
0.90	1.0	1.20	2.0	5.0	0.02	

**Table A2.** Parameters of WT3T1 module

**表 A2.** WT3T1 模块的参数

#1	#2	#3	#4	#5	#6	#7	#8
1.25	4.95	0.0	0.7e-02	21.98	0.0	1.8	1.5

**Table A3.** Parameters of WT3P1 module

**表 A3.** WT3P1 模块的参数

#1	#2	#3	#4	#5	#6	#7	#8	#9
0.30	150.0	25.0	3.0	30.0	0.0	27.0	10.0	1.0

**Table A4.** Parameters of WT3E1 module

**表 A4.** WT3E1 模块的参数

#1	#2	#3	#4	#5	#6	#7	#8
0.15	18.0	5.0	0.0	0.05	3.0	0.60	1.12
#9	#10	#11	#12	#13	#14	#15	#16
0.10	0.296	-0.436	1.10	0.05	0.45	-0.45	5.0
#17	#18	#19	#20	#21	#22	#23	#24
0.05	0.90	1.20	40.0	-0.50	0.40	0.05	0.05
#25	#26	#27	#28	#29	#30	#31	
1.0	0.69	0.78	0.98	1.12	0.74	1.20	

**Table A5.** Parameters of REGCAU1 module

**表 A5.** REGCAU1 模块的参数

#1	#2	#3	#4	#5	#6	#7
0.2e-01	10.0	0.90	0.50	1.22	1.20	0.80
#8	#9	#10	#11	#12	#13	#14
0.40	-1.30	0.2e-01	0.70	9999.0	-9999.0	1.0

**Table A6.** Parameters of REECBU1 module**表 A6.** REECBU1 模块的参数

#1	#2	#3	#4	#5	#6	#7	#8	#9
-99.0	99.0	0.0	-0.5e-01	0.5e-01	0.0	1.05	-1.05	0.0
#10	#11	#12	#13	#14	#15	#16	#17	#18
0.5e-01	0.436	-0.436	1.10	0.90	0.0	0.10	0.0	40.0
#19	#20	#21	#22	#23	#24	#25		
0.2e-01	99.0	-99.0	1.0	0.0	1.82	0.2e-01		

**Table A7.** Parameters of REPCAU1 module**表 A7.** REPCAU1 模块的参数

#1	#2	#3	#4	#5	#6	#7	#8	#9
0.2e-01	18.0	5.0	0.0	0.75e-01	0.0	0.0	0.0	0.2e-01
#10	#11	#12	#13	#14	#15	#16	#17	#18
0.10	-0.10	0.0	0.0	0.436	-0.436	0.10	0.5e-01	0.25
#19	#20	#21	#22	#23	#24	#25	#26	#27
0.0	0.0	999.0	-999.0	999.0	-999.0	0.10	20.0	0.0