

6 课程报告

基于深度学习的台风强度遥感分类与预测

摘要: 台风作为最具破坏性的自然灾害之一，其精准分类与预测对防灾减灾具有重要意义。传统台风监测方法依赖气象站数据和 Dvorak 分析方法进行台风定级，存在流程复杂、专家经验依赖性强等局限性。本文基于日本国立情报研究所提供的台风云图数据，构建了 ResNet18、VisionTransformer 和 VGG 三种深度学习分类模型。实验表明，ResNet18 模型表现最优，分类准确率达 93%，而 VisionTransformer 和 VGG 分别仅获得 66.2%和 53.4%的准确率。为进一步提升性能，提出两种改进模型：1)融合 Dvorak 物理特征的 ResNet-Dvorak 模型，分类精度提升至 93.2%；2)引入 LSTM 的 ResNet-LSTM 时序模型，但因数据量不足出现明显过拟合（测试集准确率下降至 61%）。研究表明，传统 CNN 架构在台风静态图像分类任务中更具鲁棒性，而物理特征融合可有效提升模型泛化能力。

关键词: 台风；云图；深度学习；Dvorak；强度分类；预测；ResNet；LSTM

1. 数据与方法

1.1 数据描述

本实验采用的数据来源于日本国立情报研究所(National Institute of Informatics, NII)，原始数据来自于“Himawari”系列卫星，使用了 2005 至 2023 年 203 个台风共 22338 张图像，考虑到不同年代的图像波段组成不同，IR1 是 Himawari 卫星自 1978 年以来唯一持续可用的通道，而可见光早期噪声大，因此全部数据采用红外单一波段，时间分辨率为 1 小时，即每隔一小时对系列台风即有一个成像，空间分辨率为 5km。

而台风的标签来源于日本气象厅，不同的国家对台风等级的划分具有不同的标准，日本气象厅则采用世界气象组织建议的十分钟平均风速作为持续风速来区分热带气旋等级。

表 1 日本气象厅台风分级

分类等级	台风等级	最大风速/kt
1	热带低气压(Tropical depression,TD)	[0,34)
2	热带风暴(Tropical storm,TS)	[34,64)
3	强台风(Strong typhoon,STTY)	[64,85)
4	超强台风(Super typhoon,SUTY)	[85,105)
5	猛烈台风(Violent typhioon,VITY)	[105,∞)

数据范围覆盖了东经 100 度至东经 180 度的北太平洋西部地区。包括海上台风以及登录台风。

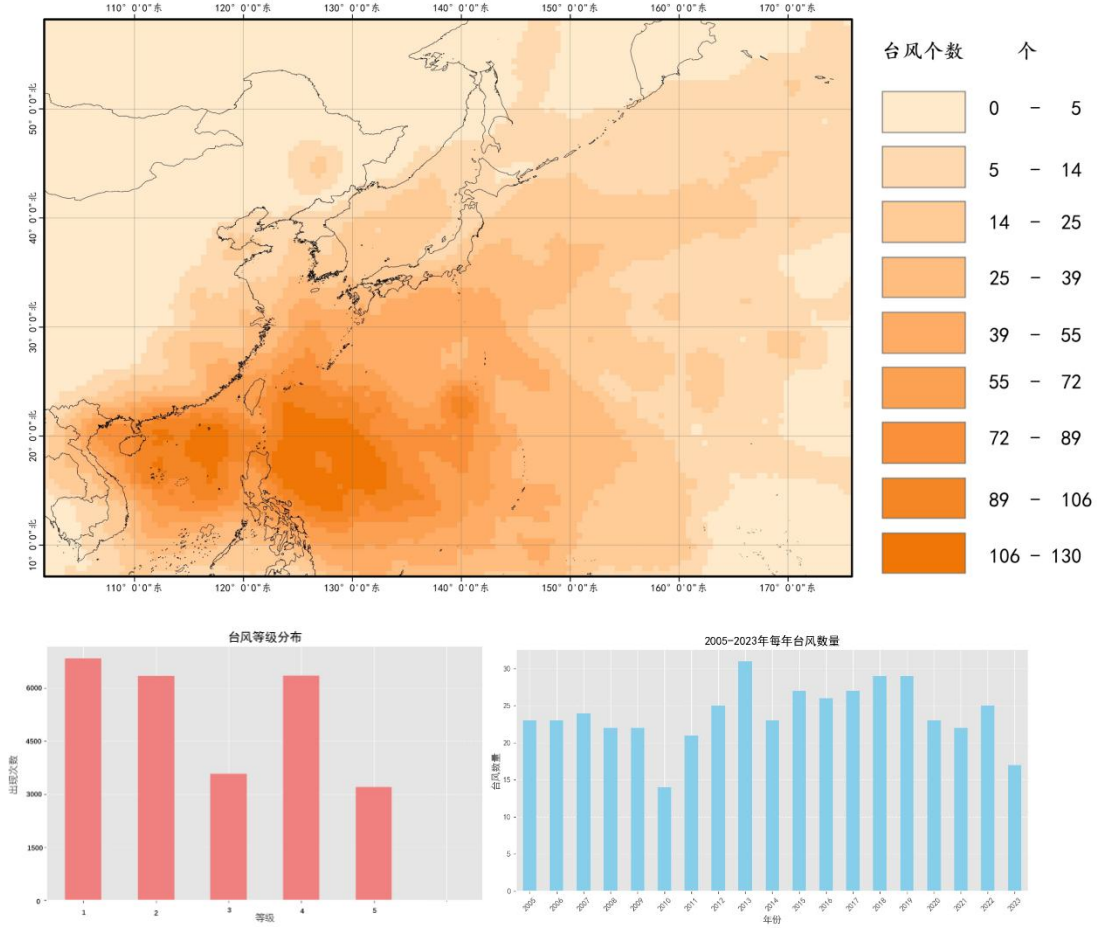


图 1 台风数据概览统计

图 1 展示了所用数据集的基本情况。从图中可以明显看出，台风活动主要集中在西北太平洋海域，尤其是菲律宾以东、南海北部至我国东南沿海一带，台风路径点密度显著高于其他区域，最高可达每格栅格 130 个路径点。另外，1、2、4 级数量均匀，3、5 级数量较少，年分布数总体来说较均匀。

1.2 数据预处理

为消除不同卫星传感器间的辐射量纲差异并适配深度学习模型的输入要求，同时保证梯度下降更稳定，本研究对原始红外亮温数据，本实验采样进行线性归一化处理。

$$T_{\text{norm}} = \frac{T - T_{\text{max}}}{T_{\text{max}} - T_{\text{min}}}$$

然后使用 pytorch 的工具进行降采样以降低计算成本，经过降采样后，图像大小由 512×512 降低为 224×224 ，其具体的降采样方法为：

首先进行两次水平方向插值：

$$f(x, y_1) = \frac{x_2 - x}{x_2 - x_1} f(x_1, y_1) + \frac{x - x_1}{x_2 - x_1} f(x_2, y_1)$$

$$f(x, y_2) = \frac{x_2 - x}{x_2 - x_1} f(x_1, y_2) + \frac{x - x_1}{x_2 - x_1} f(x_2, y_2)$$

然后进行一次垂直方向插值：

$$f(x, y) = \frac{y_2 - y}{y_2 - y_1} f(x, y_1) + \frac{y - y_1}{y_2 - y_1} f(x, y_2)$$

由于台风具有南北差异，所以在数据，在台风的风强分类过程中，模型可能学习到了其螺旋特征，所以在数据增强过程中不宜对图像进行旋转处理。

1.3 模型与方法

实验使用卷积神经网络（CNN）通过卷积操作自动提取局部特征，实现对输入数据的高层次表达。为探究不同模型的表现，本文使用了 ResNet、VGG 和 Vision Transformer 三个经典分类模型进行对比。

深层 CNN 容易出现梯度消失等问题，影响训练效果。而残差网络（ResNet），可以通过残差连接缓解梯度消失，提升网络深度和性能。具体地，残差块将输入与多层变换后的特征相加，实现信息的直接传递。即：

$$\mathbf{y} = \mathcal{F}(\mathbf{x}, \{W_i\}) + \mathbf{x}$$

其中， \mathbf{x} 为输入特征， $\mathcal{F}(\mathbf{x}, \{W_i\})$ 表示残差块中的卷积层及激活函数构成的残差映射， \mathbf{y} 为残差块的输出。

VGG 网络结构相对简单，由多个连续的小卷积核叠加构成，通过堆叠卷积层来增加网络深度，从而增强模型的非线性表达能力。尽管参数量较大，但其结构清晰，易于移植和改进。

Vision Transformer（ViT）则采用完全不同于传统 CNN 的架构。ViT 利用自注意力机制从全局角度建模输入图像中各区域之间的长距离依赖关系，尤其适合处理如台风图像中具有复杂空间结构的对象。其核心思想是将整张图像划分为若干固定大小的图像块（patch），并将每个 patch 展开后嵌入为向量序列，输入至标准的 Transformer 编码器中进行处理。

为适应红外遥感图像的特征提取需求，本文以 ResNet18 为主干网络，引入 DvorakConv 模块，构建了改进模型 ResNet-Dvorak，如图所示：

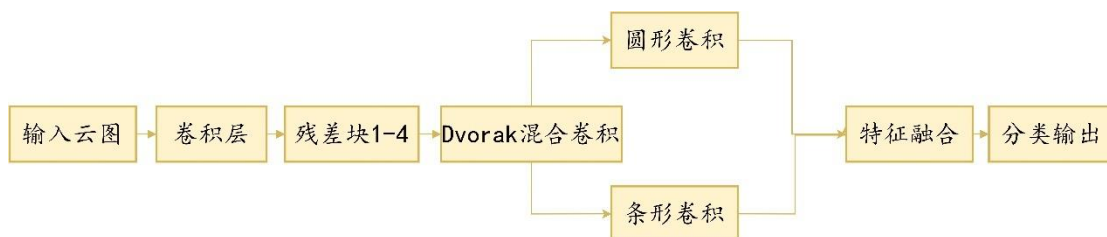


图 2 ResNet-Dvorak 模型结构

具体而言，模型以单通道红外图像作为输入，初始卷积层将输入映射到 64 维特征空间，并经过批归一化、ReLU 激活与最大池化操作后，依次通过 ResNet18 中四个标准残差层（Layer1~Layer4）提取多层次深度特征，然后在此基础上额外加入 DvorakConv 模块，用以模拟气象学中“德沃夏克分析法”中强调的圆形云团结构与带状云带结构。

DvorakConv 模块由两个分支组成：一个是标准 3×3 卷积，用于提取中心对称的环状云系特征，另外一个为两个串联的方向性卷积（ 3×1 与 1×3 ），模拟带状卷积结构以捕捉条带状云带特征。两个分支的输出特征沿通道维度拼接后，经过一个 1×1 卷积融合，并接 ReLU 激活，进一步压缩维度并增强表达能力。最终，通过自适应全局平均池化压缩为 1×1 特征向量，输入至全连接层输出台风强度预测结果

为充分利用卫星图像中的时序变化信息，本文还将 ResNet 与长短期记忆网络（LSTM）结合构建 ResNet-LSTM 实现时间序列的云图分类和预测。模型首先去除了 ResNet 的分类层，仅保留其特征提取部分，即将全局平均池化层提取到的空间特征序列（形状为 (B, T, F) ，其中 B 为批次大小， T 为时间步， F 为特征维度）输入到一个 LSTM 模型中。模型每一步输出一个分类值和预测值。

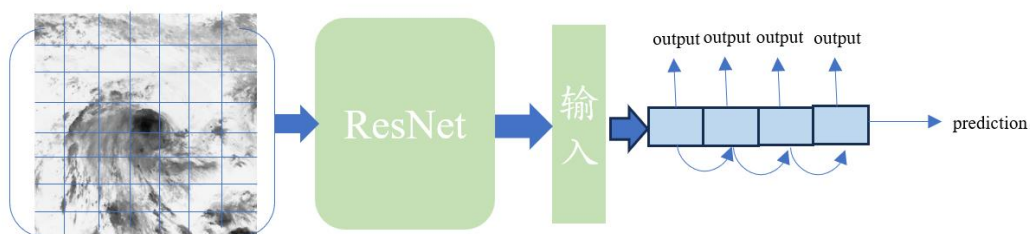


图 3 ResNet-LSTM 模型结构

2. 实验与结果

2.1 实验设置

本研究在 Ubuntu 20.04 LTS 操作系统环境下完成模型的训练与评估，ResNet、VGG、ViT 模型的训练使用搭载 NVIDIA RTX 3050 Laptop（4GB 显存）的笔记本电脑，一次训练大约 3 个小时。ResNetLSTM 使用搭载 NVIDIA GTX 1080 Ti(11GB 显存)的云服务器进行训练，一次训练大约 8 小时。统一使用深度学习框架 PyTorch 2.1。

表 2 超参数设置

参数名称	设置值
优化器	Adam
学习率	0.0001
Batch Size	16
训练轮次 (Epochs)	100
数据增强	随机裁剪、亮度扰动
早停策略	EarlyStopping (Patience=10)
滑动窗口策略	StepSize=4
分类模型损失函数	CrossEntropyLoss
预测模型损失函数	MSELoss

3.2 结果分析

(1) 基础模型分类结果

图 4 展示了 ResNet、VGG 和 ViT 三种模型在训练过程中的损失函数变化趋势，可以明显看出三者收敛速度与最终损失值上存在较大差异。其中，ResNet 模型在初始阶段损失迅速下降，并在 10 轮左右趋于收敛，最终损失接近于 0，表现出良好的优化效果。ViT 模型的损失下降相对平缓，最终也能达到较低水平；而 VGG 模型损失虽快速下降，但收敛后仍维持在较高值，说明其拟合能力较弱。

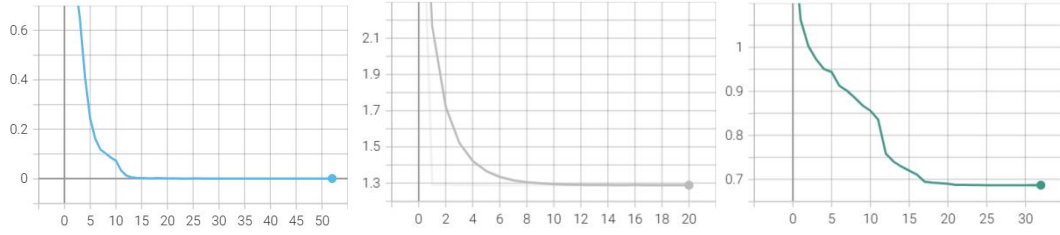


图 4 ResNet、VGG、ViT 训练时损失变化

从分类准确率指标来看，ResNet 在训练集和测试集上分别达到 99.5% 和 93.2%，表现最为优异，说明其深层残差结构在特征提取方面具有明显优势，并具备较强的泛化能力。ViT 模型的训练准确率为 73.8%，测试准确率为 66.2%，虽然低于 ResNet，但在视觉 Transformer 结构的辅助下，仍表现出一定的全局建模能力，适用于具有复杂空间关系的任务。相比之下，VGG 模型的准确率明显偏低，在训练集和测试集上分别仅为 53.8% 和 53.4%，表明其浅层结构难以充分提取复杂特征。

综上所述，三种模型中 ResNet 整体性能最优，ViT 次之，VGG 则存在较大性能瓶颈，后续工作可基于 ResNet 或 ViT 进行进一步改进与调优。

表 3 各模型准确率

训练模型	训练集准确率	测试集准确率
ResNet	0.995	0.932
VGG	0.538	0.534
ViT	0.738	0.662

(2) ResNet-Dvorak 模型结果

从训练和验证准确率曲线可以看出，ResNet-Dvorak 模型在约第 25 轮后迅速趋于收敛，训练集准确率接近 100%，验证集准确率稳定在 94.5%左右，说明模型在无过拟合迹象的前提下成功学习了数据中的判别特征。

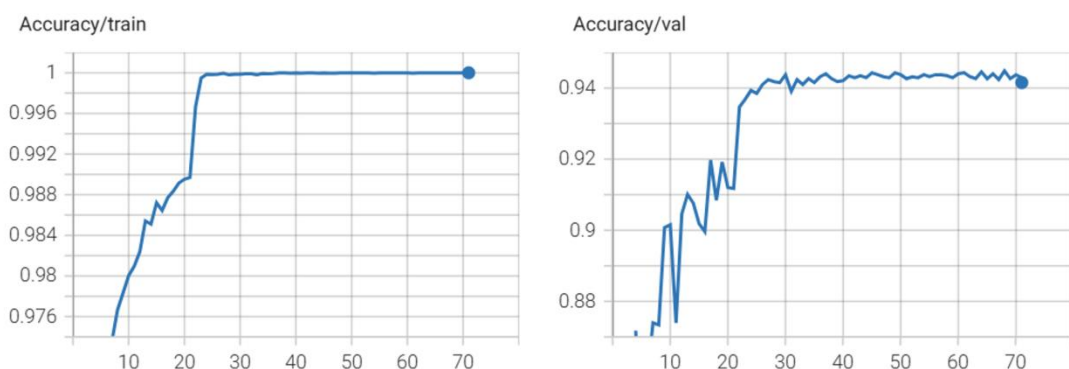


图 5 ResNet-Dvorak 训练和验证正确率

从测试集的混淆矩阵来看，第2类与第5类的识别准确率分别达到了 97.77% 与 97.51%，表现最为稳定；第1类准确率相对较低，仅为 70.95%，且被误判为第2类的比例高达 17.32%，说明这两类样本在特征空间中存在一定重叠，模型较难区分。此外，第3类的识别准确率为 87.16%，部分被误分为第1类和第4类，表明该类别在边界特征上存在模糊。

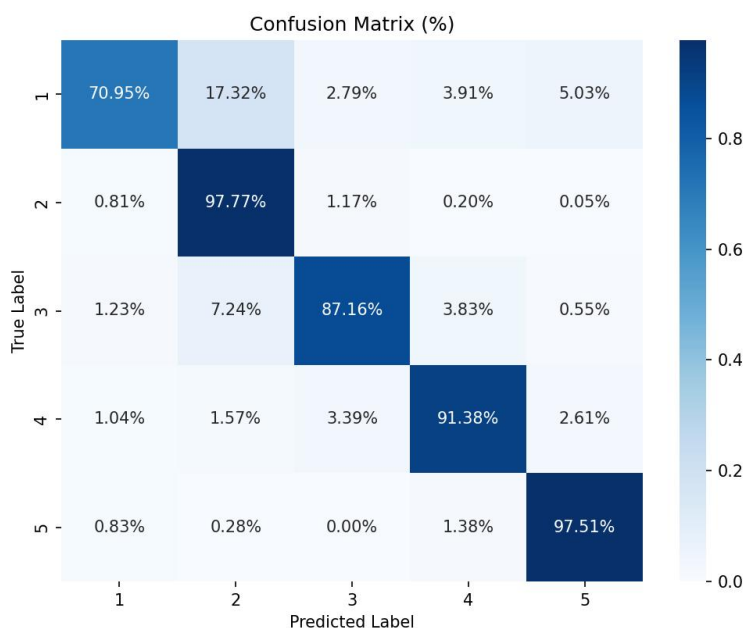


图 5 ResNet-Dvorak 测试集的混淆矩阵

利用 Grad-CAM 热力图可视化模型，图 6 说明模型能够将关注点集中在台风的中心区域，并以圆形扩展区域。展示出在台风强度较低（1、2 级）时，模型更多地关注于核心区的整体特征；而在高等级（4、5 级）样本中，模型的关注区域扩大。

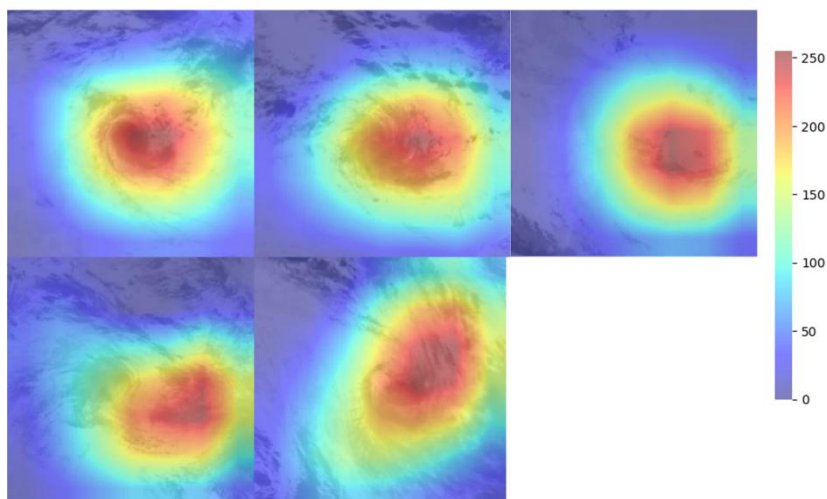


图 6 Grad-CAM 可视化结果（从左到右，从上到下依次为 1、2、3、4、5 级）

（3）ResNet-LSTM 分类模型训练结果

从训练损失曲线可以看出，随着训练轮数的增加，训练损失迅速下降并趋于收敛，表明模型能够较好地拟合训练数据。然而，验证损失则逐渐上升，且波动较大，显示出明显的过拟合现象。这说明模型在训练集上的学习效果虽好，但在验证集上的泛化能力有限。

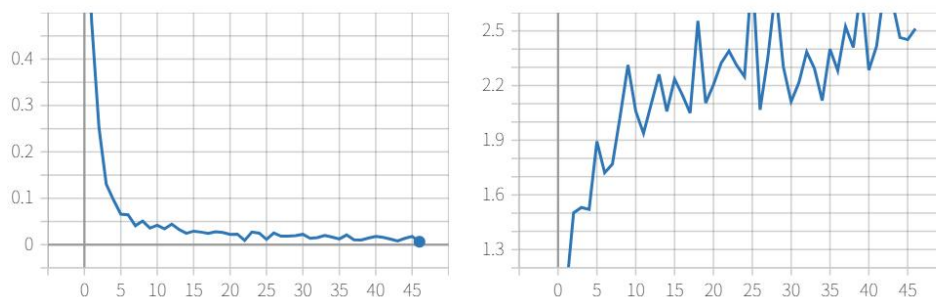


图 7 ResNet-LSTM 训练（左）、验证（右）损失变化

从验证集上的分类精度变化也可以看到，模型验证精度整体波动较大，最高点约为 0.63，最低点接近 0.57，整体数值较为不稳定，没有出现明确的上升趋势。这同样说明模型存在过拟合的问题。在最终的测试集上，模型精度为 0.62，效果较差。

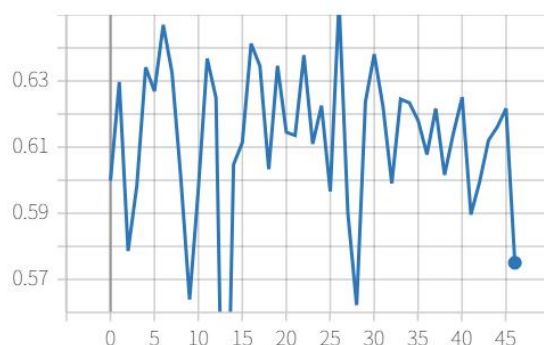


图 8 ResNet-LSTM 验证集分类精度、

(4) ResNet-LSTM 预测模型训练结果

和 ResNet-LSTM 分类模型一样，预测模型也出现了过拟合现象，从左图可以看到，随着训练轮数的增加，训练损失持续快速下降，并最终趋于较低且稳定的水平，说明模型在训练集上表现良好，能够较好地拟合训练数据。

然而，右图显示，验证损失整体波动较大，并未出现随训练进程而显著下降的趋势，且在某些阶段存在较高的峰值，这表明模型在验证集上的泛化能力有限。尤其在训练后期，虽然训练集损失已较低，但验证损失曲线依然起伏较大，部分阶段甚至有逐步上升趋势，反映出一定的过拟合现象。

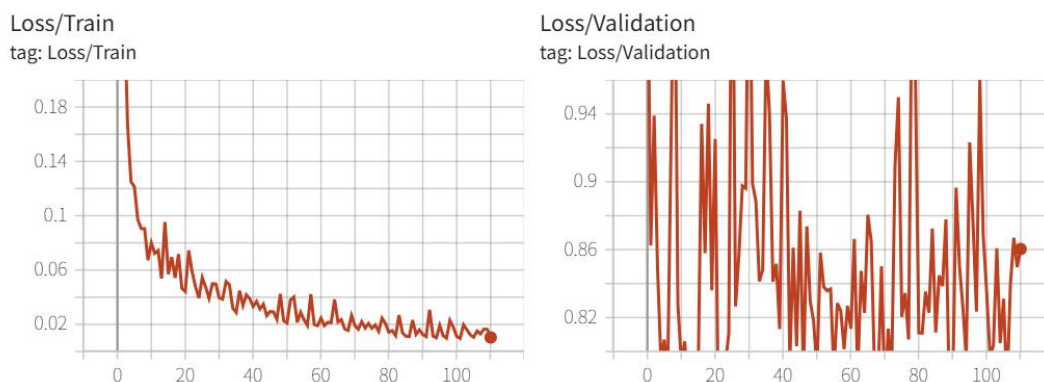


图 9 ResNet-LSTM 预测模型损失变化

3. 讨论与展望

本研究围绕台风强度识别与趋势预测任务，构建并对比了多种深度学习模型，包括 ResNet、VGG、ViT、ResNet-Dvorak 和 ResNet-LSTM 等，系统评估了各类模型在任务场景中的适应能力与表现差异。

首先，从基础分类模型的对比结果来看，ResNet 的特征提取能力和泛化性能方面明显优于其他模型，其在图像识别上体现了较强的稳定性和准确性。ViT 模型尽管未达到 ResNet 的准确率，但通过引入 Transformer 结构，展现出一定

的全局感知能力，未来可通过改进训练策略进一步提升性能。相比之下，VGG 模型因结构较浅，但是参数很多，在复杂图像场景下难以充分提取判别特征，导致准确率过低。

其次，基于 Dvorak 分析法设计的 ResNet-Dvorak 模型在分类准确率和可解释性方面均表现出色，不仅能精确区分大多数等级的台风，通过 Grad-CAM 可视化发现聚焦于台风中心区域，且呈现圆形区域，说明模型学习到了台风分级特征，具备较强的应用潜力。

然而，在引入时序信息的 ResNet-LSTM 分类与预测模型中，均出现明显的过拟合现象。尽管训练损失迅速下降，但验证集损失波动较大且存在反弹。说明模型可能存在训练时数据量不足、模型参数量过大、ResNet 提取的空间特征与 LSTM 的时序建模未充分融合等问题。

虽然在实验过程中我尝试使用更多的数据，但是受于时间和算力限制，未产生更好的效果，未来可能在此基础上进一步融合 ResNet-LSTM 模型，引入 Dropout 和正则化机制，抑制过拟合的发生。

参考文献

- [1] KITAMOTO A, HWANG J, VUILLOD B, et al. Digital Typhoon: Long-term satellite image dataset for the spatio-temporal modeling of tropical cyclones[C]// NeurIPS 2023 Datasets and Benchmarks (Spotlight).
- [2] KITAMOTO A, DZIK E, FAURE G. Machine Learning for the Digital Typhoon Dataset: Extensions to Multiple Basins and New Developments in Representations and Tasks[EB/OL]. arXiv: 2411.16421, 2024.
- [3] 2023.蔡振. 基于改进 Swin Transformer 的台风遥感图像分类方法研究 [D]. 甘肃: 西北师范大学, 2024.
- [4] 侯慧,吴文杰,魏瑞增,何浣,王磊,李正天,林湘宁.基于注意力机制的 CNN-LSTM-XGBoost 台风暴雨电力气象混合预测模型[J].智慧电力,2024,52(10):96-102