

基于改进ConvNeXt的接触网吊弦故障诊断算法的研究

李 睿, 邓晓军*, 曾 卫

湖南工业大学计算机学院, 湖南 株洲

收稿日期: 2024年3月25日; 录用日期: 2024年4月22日; 发布日期: 2024年4月30日

摘要

随着高铁运行速度与频次的增加, 接触网吊弦故障的发生频率也随之增加, 这不仅给高铁运行安全带来巨大隐患, 同时也对铁路维修工作提出了严峻挑战。现有的检测方法难以满足对吊弦故障高效准确诊断的需求, 鉴于此现状, 本文提出了一种创新的接触网吊弦故障诊断方法——InCANE Xt网络模型。该模型由ConvNeXt模型基础上改进而来, 主要改进措施包括: 通过减少模型基础块的堆叠数以降低模型复杂度, 引入Inception结构以优化卷积计算过程, 并在模型中集成了CA与SimAM注意力机制以增强特征提取能力。上述优化措施使检测模型能更准确、快速地检测吊弦的运行状态。将所提出方法与多种图像分类算法在真实接触网吊弦图像数据集上进行对比实验, InCANE Xt模型在吊弦故障诊断任务上达到了88.71%的准确率。实验结果证明, 所提出的方法在吊弦故障检测任务上展现了卓越的性能。

关键词

高速铁路, 智能检测, 深度学习, 机器视觉, 故障诊断

Research on High-Speed Railway Catenary Dropper Fault Diagnosis Based on Improved ConvNeXt

Rui Li, Xiaojun Deng*, Wei Zeng

School of Computer Science, Hunan University of Technology, Zhuzhou Hunan

Received: Mar. 25th, 2024; accepted: Apr. 22nd, 2024; published: Apr. 30th, 2024

*通讯作者。

Abstract

As the operating speed and frequency of high-speed trains increase, the incidence of catenary dropper failures also rises, posing significant safety hazards to high-speed rail operation and presenting severe challenges to railway maintenance work. Existing detection methods struggle to meet the demand for efficient and accurate diagnosis of dropper faults. In light of this situation, this paper proposes an innovative method for diagnosing catenary dropper faults—the InCANeXt network model. This model is an improvement on the ConvNeXt model, with major enhancements including reducing the number of basic blocks stacked to decrease model complexity, incorporating the inception structure to optimize the convolution process, and integrating CA and SimAM attention mechanisms to enhance feature extraction capability. These optimizations enable the detection model to diagnose the operational status of droppers more accurately and swiftly. By comparing the proposed method with various image classification algorithms on a real catenary dropper image dataset, the InCANeXt model achieved an accuracy of 88.71% in the task of dropper fault diagnosis. The experimental results demonstrate that the proposed method exhibits outstanding performance in the task of dropper fault detection.

Keywords

High-Speed Railways, Intelligent Inspection, Deep Learning, Machine Vision, Fault Diagnosis

Copyright © 2024 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

接触网系统是在电气化铁路中架设在铁路沿线上方给电力机车提供电能的大型输电系统，是高速铁路的关键子系统，是维持高速列车持续可靠供电的关键设备[1]。该系统由承力索、吊弦和接触线等部件组成，结构如图1所示。接触网对于保障列车运营安全至关重要，一旦其发生故障，轻则导致列车延误，重则损坏弓网系统，从而演变成重大安全事故[2]。

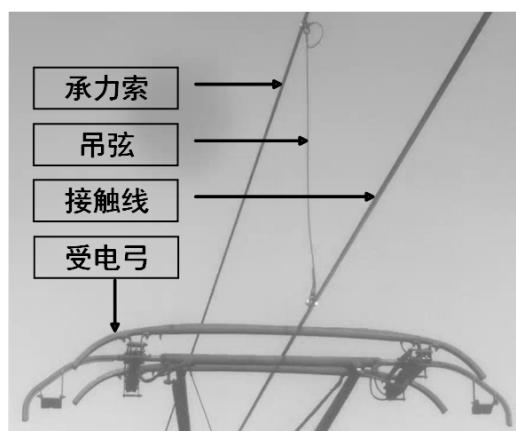


Figure 1. Catenary structure

图 1. 接触网示意图

而吊弦是高铁接触网装置中的重要组成部分，数量最多、故障概率最高，其工作状态的正常与否极大的影响到了接触网的运行状态，及时发现、处理故障的吊弦对提升高铁系统的安全性有重大意义。随着高铁的运行，受电弓会与接触网不断碰撞、摩擦，导致吊弦容易发生松弛、断裂等故障。而吊弦发生故障将会严重影响高速铁路弓网关系以及受电弓的受流质量，从而影响高铁的安全运行[3] [4] [5]。

因此，在接触网故障检测任务中，对吊弦进行详细的故障检测至关重要，是保障高铁安全运行的重要环节。然而，尚未有成熟的智能检测系统被应用于接触网吊弦的检测工作中。目前，这项工作主要还是通过人工观看离线吊弦图像的方式，甄别出吊弦存在安全隐患的地方。但是，由于视觉疲劳、精力有限等因素，再加之接触网吊弦图片比较大，吊弦占整幅图像比例比较小，以及存在的过度曝光、对比度较低等问题，人工检测很难对大量吊弦图像进行高效可靠的检测。

而自 2012 年以来，基于深度学习的方法在图像分类、目标检测与缺陷检测任务中表现出了巨大优势[6]。涌现了许多专门用于故障诊断的深度学习模型。例如 Li 提出了一种改进后的 YOLOv5 模型 YOLOv5s-D，在主干网络中加入了 SK 注意力模块，实现了对小目标特征的提取。将提取到的特征输入双向特征金字塔中进行多尺度的特征融合；并将 YOLO 头替换为了解耦头提高了模型的收敛速度与检测精度。最终在吊弦数据集上平均准确率达到了 92.9%，并且微小缺陷的检测准确率达到了 79.2%，与 YOLOv5 相比提升明显[7]。而 Tan 使用了一种基于 MaskR-CNN 算法用于接触网的智能缺陷检测，在真实图像数据中获得了 99.0% 的召回率与 92.5% 的精度，最终实现了高精度的接触网绝缘子的缺陷检测[8]，但该方法存在检测速度较慢的问题，无法满足实时检测的需求。还有研究者将 YOLO 模型与 RCNN 模型结合起来用于吊弦的故障检测[9]，先使用轻量级 YOLOv3 网络来提取吊弦区域，再使用 FasterR-CNN 来检查吊弦上的异物与弯曲，通过实验表明，FasterR-CNN 对吊弦区域提取后的异物检查与弯曲检查中取得了 90% 以上的准确率。但是依旧存在处理速度较慢的问题，而且不经其他处理的轻量化 YOLOv3 模型识别小目标的能力有限，误差会累积至后续的异物检查与弯曲检查中，影响检查的准确率。

而本文受此启发，先使用了一个目标检测模型进行吊弦的定位，获取了吊弦图像。随后设计了一种改进 ConvNeXt 的图像分类模型用于吊弦故障类型的分类，该模型以 ConvNeXt 作为基底模型，引入 Inception 结构，分解基础块中的 7×7 的大卷积核为一个 3×3 ，一个 9×1 与一个 1×9 的卷积核，并引入协调注意力机制(Coordinate Attention, CA)与无参数注意力机制(Simplified Attention Module, SimAM)，使模型在不引入更多参数的情况下，能捕获输入图片的深层次特征。本文提出的 InCANeXt (Incption-Coordinate Attention ConvNeXt) 模型与 AlexNet [6]，ResNet [10] 与 MobileNetV2 [11] 模型进行对比。实验结果表明，本文提出的方法在检测进度上有更好的性能。

2. 接触网吊弦故障诊断模型

2.1. ConvNeXt 图像分类模型

ConvNeXt [12] 模型是一种先进的图像分类模型，其结合了传统 CNN 的强大特征提取能力和 Transformer 的全局感知能力来提升图像分类的性能。它对 ResNet50 神经网络架构进行了现代化的重构，采用了类似于 Transformer 的层次结构，同时维持了卷积层的核心特性。ConvNeXt 模型结构如图 2 所示。ConvNeXt 模型改善了层次堆叠，将原 ResNet50 中四个层次的堆叠次数从(3,4,6,3)调整为了(3,3,9,3)与 SwinTransformer 的堆叠次数一致。调整后，准确率有 0.6% 的提升。并改善了下采样模块，使用一个卷积核大小为 4×4 ，步距为 4 的卷积层作为下采样层，替换了 ResNet50 中的大核卷积与最大池化下采样层，降低模型参数，并略微提升准确率。在基础块中将传统卷积替换为了深度可分离卷积，并大胆的采用了 7×7 的大卷积核作为深度可分离卷积层的卷积核。在这一系列的优化下，最终 ConvNeXt 模型在相近 FLOPs 的情况下，在 ImageNet 22K 数据集下分类准确率超过了 SwinTransformer [13]，达到了 87.8%，这

一结果暗示了在图像处理任务中，自注意力机制(Self-Attention)并不一定比卷积更有优势，为图像处理领域的模型设计提供了新的视角和思路。

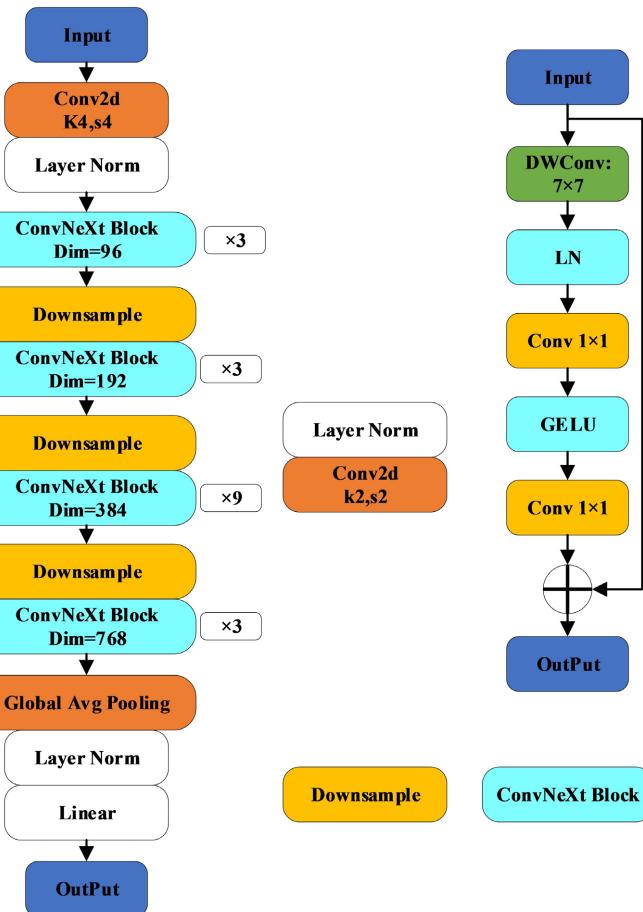


Figure 2. ConvNeXt structure
图 2. ConvNeXt 结构图

2.2. Inception 结构

Inception 结构最早被提出于 2015 年[14]，随后连续通过多年的迭代、优化，最终提出了融合 Inception 结构的 ResNet 模型[15] [16]。Inception 结构的主要思路即为将不同大小的卷积核组合在一起，与将大的卷积核分解为多个小的卷积核。将不同大小的卷积核叠在一起后，不但可以增大感受野，还可以提高神经网络的鲁棒性。而将大的卷积核分解为多个小卷积核后，可以提升运算速度，比如将 5×5 的卷积核分解为两个 3×3 的卷积核，计算开销降低了 28%。

2.3. CA 注意力机制

CA 注意力机制由 Hou 在 2021 年提出[17]，Hou 认为尽管现有的 CBAM 注意力[18]、SE 注意力[19]等注意力机制可以提升模型的性能，但上述注意力机制在计算通道注意力时通常采用全局最大池化或平均池化，而忽略了位置信息，而这种做法会影响空间注意力图的生成。因此，作者期望在引入通道注意力机制的同时，与空间注意力机制相结合，CA 注意力机制的结构如图 3 所示。

首先分别在水平坐标方向和垂直坐标方向对输入的特征进行全局平均池化，随后对每个通道进行独立

编码。经平均池化后，特征在垂直和水平方向上分别被映射，形成 $[C, H, 1]$ 和 $[C, 1, W]$ 的特征图。上述变换分别沿着两个空间方向聚合特征，从而产生一对方向感知的特征图，这允许注意力块捕获沿一个空间方向的长程依赖，同时保存另一个空间方向的精确位置信息。这种设计有助于网络更准确地定位感兴趣的目标。

随后，将水平方向的特征图转置，然后与垂直方向的特征图进行堆叠，此时得到的特征图形状为 $[C, 1, H + W]$ ，这时通过卷积、池化和激活函数调整特征图的尺寸后获得中间特征图。之后再次将特征图分离为 $[C, 1, H]$ 和 $[C, 1, W]$ 的特征图。最后，通过一个 1×1 的卷积调整通道数，并使用 sigmoid 激活函数分别获得两个维度上的注意力特征。最后，将这些注意力特征与原始特征相乘，即可获得最终 CA 注意力输出的特征图。

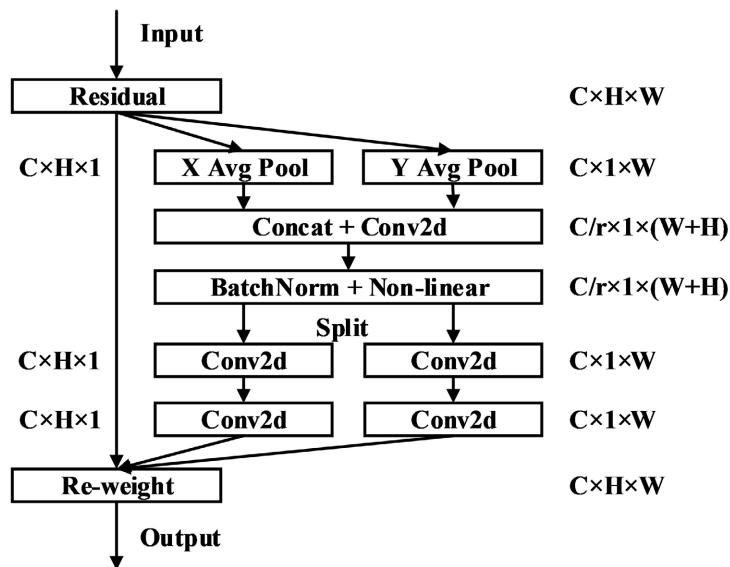


Figure 3. CA attention structure
图 3. CA 注意力机制示意图

2.4. SimAM 注意力机制

SimAM 注意力机制由 Yang 等人于 2021 年提出[20]。SimAM 注意力机制的结构如图 4(c)所示。

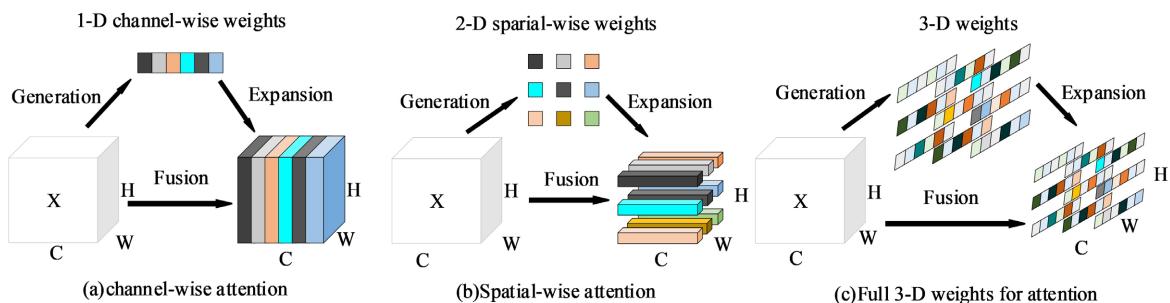


Figure 4. Attention structures. Where (a) is channel attention; (b) is spatial attention; (c) is SimAM attention.

图 4. 注意力示意图，其中(a)为通道注意力示意图；(b)为空间注意力示意图；(c)为 SimAM 注意力机制示意图

作者认为形式如图 4(a)的通道注意力模块只注重通道间的联系，而不注重位置间的信息；而形如图 4(b)的空间注意力机制则只注重位置间的信息，不注重通道间的联系。而在人类思考过程中，这两种机制

是共同存在的。因此，作者提出了 SimAM 注意力机制来模仿类似的操作。而为了评估每个神经元的重要性，作者利用了在视觉神经科学中存在的空间抑制现象，认为显示出明显的空间抑制效应的神经元在视觉处理中应该被给予更高的优先级。找到这些神经元最简单的实现是测量一个目标神经元和其他神经元之间的线性可分离性。基于此发现，作者定义了每个神经元的能量函数为式(1)：

$$e_t(\omega_t, b_t, y, x_i) = (y_t - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_o - \hat{x}_i)^2 \quad (1)$$

其中， $\hat{t} = \omega_t t + b_t$ ， $\hat{x}_i = \omega_i x_i + b_i$ ，分别为 t 与 x_i 的线性变换，其中 t 与 x_i 是输入特征 $X \in \mathbb{R}^{C \times H \times W}$ 中的单个通道中的目标神经元与其他神经元； i 是空间维度上的索引； $M = H \times W$ 是该通道上神经元的数量； ω_t 与 b_t 是线性变换的权重与偏置。公式(1)中所有值都是标量。当 $t = y_t$ 且其他 x_i 均为 y_o 时，公式(1)达到最小，其中 y_t 与 y_o 不相同。

通过最小化该方程，相当于找到目标神经元 t 与同一通道中所有其他神经元的线性可分性。随后，为简单起见，对 y_t 与 y_o 采用二值标记，并加入正则化项。最终的能量函数如式(2)：

$$e_t(\omega_t, b_t, y, x_i) = \frac{1}{M-1} \sum_{i=1}^{M-1} (-1 - (\omega_t x_i + b_i))^2 + (1 - (\omega_t t + b_t))^2 + \lambda \omega_t^2 \quad (2)$$

理论上，每个通道存在一个能量函数。式(2)具有解析解：

$$\omega_t = -\frac{2(t - \mu_t)}{(t - \mu_t)^2 + 2\sigma_t^2 + 2\lambda} \quad (3)$$

$$b_t = -\frac{1}{2}(t + \mu_t)\omega_t \quad (4)$$

其中 $\mu_t = \frac{1}{M-1} \sum_{i=1}^{M-1} x_i$ ， $\sigma_t^2 = \frac{1}{M-1} \sum_{i=1}^{M-1} (x_i - y_i)^2$ ，是该通道除 t 外所有神经元的均值与方差。

因此，最小能量可有式(5)得到：

$$e_t^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (5)$$

其中 $\hat{\mu} = \frac{1}{M} \sum_{i=1}^M x_i$ ， $\hat{\sigma}^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \hat{\mu})^2$ 。公式(5)表明能量越低，神经元 t 与周围神经元的区别越大，重要性越高。因此每个神经元的重要性可以用 $\frac{1}{e_t^*}$ 来表示。

最终，使用 sigmoid 激活函数对特征进行增强得到式(6)：

$$\tilde{X} = \text{sigmoid}\left(\frac{1}{E}\right) \odot X \quad (6)$$

其中， E 是在通道与空间维度上能量。

2.5. InCANeXt 模型结构

受上述技术的启发，本文提出 InCANeXt 模型，模型结构如图 5 所示。先削减了 InCANeXtBlock 的数量，并使用 Inception 结构分解大卷积核，使卷积计算的开销降低了 44%。并在 InCANeXtblock 中加入 SimAM 注意力机制增强了不同层级的特征提取能力。在整个模型架构中，在每个主要阶段末尾引入 CA 注意力机制，以加深多层次特征的提取与融合，提升了模型对吊弦检测任务上的性能。

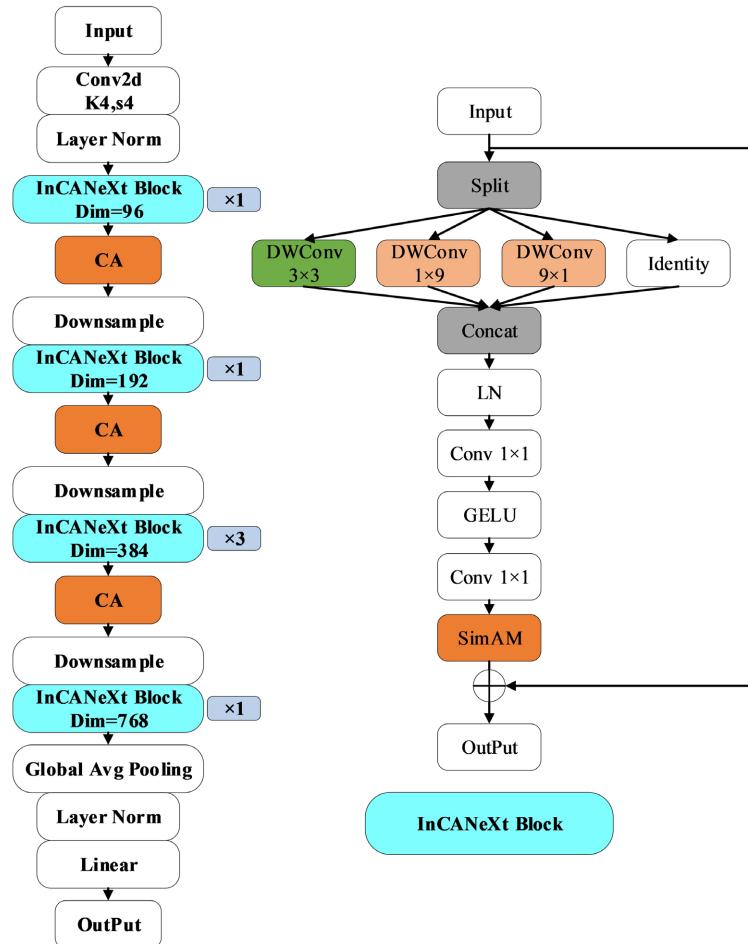
**Figure 5.** InCANeXt structure

图 5. InCANeXt 模型结构

3. 实验

3.1. 数据集

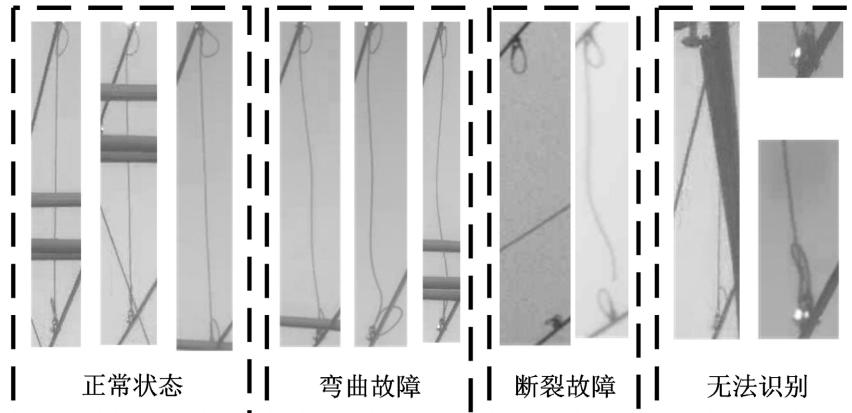
**Figure 6.** Examples of dropper: (a) normal state;(b) loose state;(c) broken state; (d) 无法识别

图 6. 吊弦图片实例: (a) 正常状态的吊弦; (b) 弯曲状态的吊弦; (c) 断裂状态的吊弦; (d) 无法识别

为了证明本文提出方法的有效性，使用真实铁路运行图像数据集作为研究对象。数据集中存在类别不平衡现象，即正常吊弦图像数量远多于故障吊弦图像，而断裂故障的图像尤其稀少。为了实现数据类别的平衡，以松弛吊弦数据量为基准，调整正常吊弦与无法识别图像的数量，使其与松弛类别图像的规模相匹配，并利用 Photoshop 等图像处理软件生成了一系列断裂吊弦图像，最终构建了一个类别较为平衡的含 2339 张吊弦图像的数据集。对吊弦数据集进行标注后，将数据集随机分为训练集、验证集与测试集，比例为 8:1:1。使用的数据集图像如图 6 所示。

3.2. 实验环境与评价指标

本次实验使用的计算机硬件平台为 Intel CoreI5-13600KF，32 GB 内存，NVIDIA RTX4070Ti，显存 12 G，操作系统为 Windows10，开发框架为 PyTorch，CUDA 版本为 11.8，python 版本为 3.8.0。

为了全面评估 InCANeXt 算法的性能，采用精度、召回率、准确率、F1 分数与 FPS 对模型性能进行评价。具体公式如(7)~(10)。

1) 精度：精度是模型正确标记为属于正类的样本的比例与真正为正类的样本总数的比值，该值越高表示模型预测能力越好，计算公式为式(7)。

$$precision = \frac{TP}{TP + FP} \quad (7)$$

2) 召回率：召回率是全部正样本中被正确预测为正样本的比例，该值越高表示模型预测能力越好，计算公式为式(8)。

$$recall = \frac{TP}{TP + FN} \quad (8)$$

其中 TP 为真阳性样本数，表示模型正确地识别出一个样本为正类。 FN 为假阴性样本数，表示模型错误地将正类样本识别为负类样本的数量。 FP 为假阳性样本数，表示模型错误地将负类样本识别为正类样本的数量。

3) 准确率：准确率为在所有样本中，预测正确的概率，其计算公式为式(9)。

$$accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (9)$$

4) F1 分数：是统计学中用来衡量二分类(或多任务二分类)模型精确度的一种指标。它同时兼顾了分类模型的准确率和召回率。 $F1$ 分数可以看作是模型准确率和召回率的一种加权平均，它的最大值是 1，最小值是 0，值越大意味着模型越好。计算公式为式(10)。

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \quad (10)$$

5) FPS：表示每秒内可以处理的图片数量，通常使用处理一张图片所需的时间来评估检测速度，时间越短，FPS 越高，速度越快。

3.3. 对比实验

为了评估所提模型的性能，在相同数据集下将所提模型与当前主流的图像分类模型在相同实验环境下进行了对比实验。实验学习率设置为 4e-3，共训练 200 个 epoch，批次大小设置为 32。对比试验结果如表 1 所示。

Table 1. Comparative experimental results
表 1. 对比实验结果

模型	精度/%	召回率/%	Top1-acc/%	F1-score/%	Fps
AlexNet	37.51	45.83	73.48	41.45	177.10
ResNet50	70.22	76.61	81.06	71.37	153.61
MobileNetV2	66.73	54.14	84.85	54.85	159.60
ConvNeXt-tiny	38.68	36.25	72.73	35.98	147.04
InCANE Xt	79.64	68.46	88.71	73.63	167.10

通过分析对比实验结果，可以看出 InCANE Xt 模型在高铁接触网吊弦故障诊断任务上的分类性能无论是在精度上还是在其他关键性能指标上均优于其他基准模型。特别值得注意的是，原始的 ConvNeXt-tiny 模型在当前分类任务中多项关键指标均处于末尾位置。对于 ConvNeXt-tiny 模型在此任务上表现不佳的现象，分析可能原因如下：ConvNeXt 模型是由 ResNet50 模型改进而来，其改进灵感部分来源于 Swin Transformer 模型，这使得 ConvNeXt 在设计之初便倾向于处理大规模数据集和复杂任务的需求。因此，当面对相对较小规模的数据集时，其复杂的结构容易引发过拟合问题，导致模型性能下降。ResNet50 与 InCANE Xt 的实验结果验证了此猜想，在一定程度上 ResNet50 与 InCANE Xt 的模型规模更适合此吊弦数据集，这也是它们能够呈现出更好性能的原因之一。

此外，可以观察到，所有参与对比实验的模型精确度都没超过 90%。分析实验结果，认为是吊弦状态的难以界定造成。如图 7 所示，可见，吊弦在轻微弯曲或局部弯曲时，分类模型会难以分辨这些吊弦是否故障。这种“临界状态”几乎贡献了全部的分类误差。因此，如何对这类界定不清的数据进行精确分类，将成为我们下一步研究的重点。



Figure 7. Refine results
图 7. 结果分析

3.4. 消融实验

为了验证本文所提模型每项改进的实际效果，进行了多组消融实验，每组实验都使用相同的训练参数。消融实验的设置与结果如表 2 所示。

Table 2. Results of ablation experiments

表 2. 消融实验结果

Inception 结构	注意力	简化	Top1-acc/%	F1-score	FPS
×	×	×	72.73	35.98	130.13
×	×	√	80.99	63.01	143.03
√	×	√	83.06	73.42	147.01
√	√	√	88.71	73.63	145.10

分析消融实验结果可知，可以看出对 ConvNeXt 模型影响最为显著的改进措施是模型参数的简化，在对 ConvNeXt 原模型进行模型简化之前，仅仅只有 72.73% 的 Top1 准确率与 35.98% 的 F1 分数，在简化后，Top1-acc 提升至 80.99%，F1-score 提升至 63.01%，而 FPS 也取得了 13 的提升。这一现象再次说明对于 ConvNeXt 这种规模较大的分类模型，面对小规模的图像数据集容易出现过拟合现象，影响分类效果。

此外，实验还发现，当在简化版本的 ConvNeXt 模型中集成 Inception 结构，将原模型中的 7×7 的大卷积核简化成一个 3×3 ，一组 1×9 与 9×1 和一个残差连接后，模型的精确度与 FPS 均出现了一定提升，最大的提升出现在 F1 分数这一指标中，说明在引入 Inception 结构后，模型对更深层次与多尺度的特征提取能力得到了增强，这符合预期中 Inception 结构优化特征提取能力的目标。

最后，通过在模型中集成注意力机制，模型的 Top1 准确率提升了 5%，这验证了 CA 注意力机制和 SimAM 注意力机制在增强模型特征提取能力方面的有效性，从而促进了整体分类性能的提高。然而，尽管注意力机制提升了模型的准确率，但其也带来了较高的计算复杂度，导致 FPS 略微下降。

4. 结论

本文提出了一种基于改进 ConvNeXt 的智能化高铁接触网吊弦缺陷检测方法。该方法针对吊弦故障诊断任务进行了优化。通过减少模型基本块减少模型复杂度，有效避免了小规模图像样本的过拟合问题。同时，集成了 Inception 结构增强了模型提取多尺度特征的能力。此外，在模型每个主要阶段融入了 CA 与 SimAM 注意力机制，提升了吊弦检测性能。最终，相较于其他图像分类模型，InCANeXt 模型取得了 88.71% 的 Top1 准确率，表明在当前接触网吊弦故障诊断任务中，InCANeXt 模型具有最优的性能。

尽管 InCANeXt 模型在吊弦诊断任务中表现出了出色的性能，但其对高性能独立显卡的依赖，限制了该模型在配置较低的环境中进行实时监测的能力。此外，如何改进数据平衡策略，通过更加科学的数据采样或增强技术改善训练集的平衡性，以减少模型对于常见类别的偏见。以及如何精确分类难以界定状态的吊弦图像，也将是未来研究的重点。

基金项目

基金项目：中国高校产学研创新基金——新一代信息技术创新项目(2022IT121)；2023 年湖南省大学生创新创业训练计划项目(S202311535056)；湖南省自然科学基金资助项目(2024JJ7148)。

参考文献

- [1] Song, Y., Wang, H., Frøseth, G., et al. (2023) Surrogate Modelling of Railway Pantograph-Catenary Interaction Using

- Deep Long-Short-Term-Memory Neural Networks. *Mechanism and Machine Theory*, **187**, Article ID: 105386. <https://doi.org/10.1016/j.mechmachtheory.2023.105386>
- [2] Yang, J., Duan, H., Li, L., et al. (2023) 1D CNN Based Detection and Localisation of Defective Droppers in Railway Catenary. *Applied Sciences*, **13**, Article 6819. <https://doi.org/10.3390/app13116819>
- [3] 袁远, 纪占玲, 陈立明, 等. 高速铁路接触网弹性吊索失效机理分析[J/OL]. 铁道科学与工程学报: 1-12. <https://doi.org/10.19713/j.cnki.43-1423/u.T20232022>, 2024-03-10.
- [4] Cho, Y.H., Lee, K., Park, Y., et al. (2010) Influence of Contact Wire Pre-Sag on the Dynamics of Pantograph-Railway Catenary. *International Journal of Mechanical Sciences*, **52**, 1471-1490. <https://doi.org/10.1016/j.ijmecsci.2010.04.002>
- [5] 李永生, 韩宝峰, 祝晓红, 等. 高速铁路接触网用锻造式无螺栓整体吊弦的设计开发[J]. 电气化铁道, 2023, 34(4): 51-56.
- [6] Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2012) ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90.
- [7] Li, Z., Rao, Z., Ding, L., et al. (2023) YOLOv5s-D: A Railway Catenary Dropper State Identification and Small Defect Detection Model. *Applied Sciences*, **13**, Article 7881. <https://doi.org/10.3390/app13137881>
- [8] Tan, P., Li, X., Ding, J., et al. (2022) Mask R-CNN and Multifeature Clustering Model for Catenary Insulator Recognition and Defect Detection. *Journal of Zhejiang University-SCIENCE A*, **23**, 745-756. <https://doi.org/10.1631/jzus.A2100494>
- [9] Cui, J., Wu, Y., Qin, Y., et al. (2020) Defect Detection for Catenary Sling Based on Image Processing and Deep Learning Method. In: Liu, B., Jia, L., Qin, Y., Liu, Z., Diao, L. and An, M., Eds., *EITRT 2019: Proceedings of the 4th International Conference on Electrical and Information Technologies for Rail Transportation (EITRT)*, Springer, Singapore, 675-683. https://doi.org/10.1007/978-981-15-2914-6_64
- [10] He, K., Zhang, X., Ren, S., et al. (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [11] Sandler, M., Howard, A., Zhu, M., et al. (2018) Mobilenetv2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 4510-4520. <https://doi.org/10.1109/CVPR.2018.00474>
- [12] Liu, Z., Mao, H., Wu, C., Y., et al. (2022) A Convnet for the 2020s. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans, 18-24 June 2022, 11966-11976. <https://doi.org/10.1109/CVPR52688.2022.01167>
- [13] Liu, Z., Lin, Y., Cao, Y., et al. (2021) Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Montreal, 10-17 October 2021, 9992-10002. <https://doi.org/10.1109/ICCV48922.2021.00986>
- [14] Szegedy, C., Liu, W., Jia, Y., et al. (2015) Going Deeper with Convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>
- [15] Szegedy, C., Vanhoucke, V., Ioffe, S., et al. (2016) Rethinking the Inception Architecture for Computer Vision. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 2818-2826. <https://doi.org/10.1109/CVPR.2016.308>
- [16] Szegedy, C., Ioffe, S., Vanhoucke, V., et al. (2017) Inception-V4, Inception-Resnet and the Impact of Residual Connections on Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, **31**, 4278-4284. <https://doi.org/10.1609/aaai.v31i1.11231>
- [17] Hou, Q., Zhou, D. and Feng, J. (2021) Coordinate Attention for Efficient Mobile Network Design. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, 20-25 June 2021, 13708-13717. <https://doi.org/10.1109/CVPR46437.2021.01350>
- [18] Woo, S., Park, J., Lee, J., Y., et al. (2018) Cbam: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018*, Springer, Cham, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1
- [19] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/CVPR.2018.00745>
- [20] Yang, L., Zhang, R.Y., Li, L., et al. (2021) SimAM: A Simple, Parameter-Free Attention Module for Convolutional Neural Networks. *Proceedings of the 38th International Conference on Machine Learning*, 18-24 July 2021, 11863-11874.