# 加速深度神经网络在硬件瞬态故障下准确性评估的端到端自动化方法

焦佳佳，闻然，杨洪
上海海事大学信息工程学院，中国上海市，201306

**摘要：**硬件瞬态故障已被证实会对深度神经网络产生显著影响，尤其在自动驾驶汽车、医疗保健和航天应用中，其安全关键性误分类概率增加多达4倍。然而，使用准确的故障注入方法进行不准确性评估非常耗时，在完整的仿真平台可能需要几个小时甚至几天时间。为加快对深度神经网络上硬件瞬态故障的评估，设计了一种统一的端到端自动化方法——A-Mean，该方法利用基本操作（如卷积、加法、乘法、激活函数、最大池化等）的静默数据损失率以及静态两级均值计算机制，快速计算整体静默数据损失率，以估算一般分类指标准确性和特定应用指标安全关键性误分类。更重要的是，采用最大策略确定深度神经网络中非顺序结构的静默数据损失边界。然后，将静态安全关键性误分类结果与一次动态无故障执行的原始数据合并，采用最坏情况方案进一步计算瞬态故障下放大的安全关键性误分类和降半的准确性。此外，以上所有步骤均已实现自动化，以便该易于使用的自动化工具可以用于快速评估多种深度神经网络上的瞬态故障。同时，定义一种新指标“故障敏感性”以表征瞬态故障导致的安全关键性误分类升高和准确率降低的变化。与最先进的故障注入方法TensorFI+在5个深度神经网络模型和4个数据集上的比较结果表明，本文提出的评估方法A-Mean实现了高达922.80倍的加速，同时其平均安全关键性误分类损失和准确率损失仅为4.20%和0.77%。A-Mean的相关结果可通过https://github.com/breatrice321/A-Mean获取。

**关键词：**分析模型；深度神经网络；硬件瞬态故障；快速评估；自动化评估工具
https://doi.org/10.1631/FITEE.2400547