用于抗噪机器故障分类的现场噪声暴露技术

Wonjun YI[1]；Jung-Woo CHOI[2]

[1,2] 韩国科学技术院电气工程学院（KAIST ），

大韩民国

摘要

对故障声音进行现场分类是机器健康监测与诊断中的一项重要课题。然而，在工厂等嘈杂环境中，机器声音总是与环境噪声混杂在一起；机器未运行时，也可能出现只有噪声的时段。因此，基于深度神经网络（DNN）的故障分类器必须能够区分噪声与机器声音，并且对混合噪声具有鲁棒性。为解决这些问题，我们研究了现场噪声暴露（ONE），即让深度神经网络模型接触在机器运行的同一环境中录制的噪声。与异常值暴露技术类似，噪声暴露通过训练深度神经网络分类器，使其针对纯噪声数据生成均匀的预测概率分布。在推理过程中，经过 ONE 训练的深度神经网络分类器将最大软最大概率作为噪声分数，并据此判断是否处于纯噪声时段。我们将 ToyADMOS2 数据集中的机器声音与噪声混合，以模拟高噪声数据。我们对采用 ONE 训练的基于残差网络的分类器进行评估，并与采用其他分布外检测技术训练的分类器进行比较。测试结果表明，让模型接触现场噪声，比使用其他噪声或检测技术能够使模型更具鲁棒性。关键词：故障分类、噪声暴露、现场噪声

1. 引言

随着智能工厂中的制造过程实现自动化，基于机器噪声的现场故障分类正受到越来越多的关注。然而，在工厂中，噪声始终存在，而机器仅在运行期间产生声音。因此，可能会出现若干个不存在机器声音、只有工厂噪声的信号时段。为了实现稳健的自主故障诊断，故障分类器需要能够适应高水平的背景噪声，并将仅含噪声的时段与机器声音区分开来。

此前，研究人员已开展分布外（OOD）检测研究，以检测位于正常数据所形成分布之外的异常值。这类技术可用于检测我们应用中的纯噪声时段。例如，在（1）中，研究人员根据由分类任务训练的深度神经网络模型预测的最大软最大概率来检测分布外样本。（2）中也研究了另一种利用自由能函数的分布外检测算法。

另一方面，异常值暴露（OE）被认为是提升分类器 OOD 性能和鲁棒性的有效方法（3, 4）。OE 使用一个与分布内数据和测试时数据互不重叠的异常值数据集。在 OE 中，模型经过训练后会对异常值数据输出均匀分布，使最大软最大概率（MSP）尽可能低。利用分布内数据与 OOD 数据的 MSP 差异，模型可以在测试时检测 OOD 样本。

本文报告了用于 OE 的噪声类型的重要性。如前所述，训练目标是：(a) 对故障类型进行分类；(b) 检测纯噪声数据。为同时实现这两个目标，我们让深度神经网络分类器接触不同类型的噪声，并比较其性能。与 OE 一样，分类器经过训练后，会为噪声信号生成均匀概率分布，并为含噪机器声音预测独热向量。不过，我们研究了用于 OE 的噪声与嵌入含噪机器声音信号中的现场工厂噪声具有相似或不同分布时的性能差异。我们将使用现场工厂噪声进行的 OE 称为现场噪声暴露（ONE），并将通过 ONE 训练的模型与接触不同类型噪声的模型进行性能比较。此外，为验证 ONE 的有效性，我们还将其与其他训练技术进行比较，例如将纯噪声数据映射为附加类别，或采用基于自由能的技术：能量分数和能量有界学习（2）。我们使用 ToyADMOS2 数据集合成含有工厂噪声的机器声音和纯噪声数据，

[1] lasscap@kaist.ac.kr[2]
jwoo@kaist.ac.kr

并对仅噪声时段检测和故障类型分类性能进行了评估。结果表明，ONE 优于其他训练技术，以及使用其他类型噪声的 OE。

2. 背景

2.1 深度神经网络分类器和 Softmax 分数

深度神经网络分类器 经过训练，用于对分布内数据进行分类 并预测 类别的 Softmax概率。分类器的训练方式是，使目标概率与预测概率之间的分类交叉熵（CCE） （目标概率由独热向量给出，预测概率 ）达到最小化。在推理时，可以通过取预测概率的最大值来计算 Softmax 分数。

$$A(x) = \mathrm{Max}(f(x)). \tag{1}$$

此处，Max(·)表示取向量最大值的运算符。在先前的研究 (1) 中，Softmax 分数被用作 OOD 的度量，即如果 Softmax 分数高于预定义阈值，则将数据视为分布内数据，否则视为 OOD 数据。

2.2 异常值暴露

对于异常值暴露，训练深度神经网络模型时，不仅使用分布内数据集 ，还使用一个用于异常值的辅助数据集 。训练深度神经网络分类器，将分布内数据 分类为真实标签独热向量 ，同时对异常值数据产生均匀分布。T总损失函数 可写为

$$\mathcal{L}_{total} = \mathbb{E}_{(x_{in},y)\sim D_{in}}(\mathcal{H}(y, f(x_{in}))) + \mathbb{E}_{x_{OE}\sim D_{OE}}(\alpha\mathcal{H}(u, f(x_{OE}))). \tag{2}$$

$\mathcal{H}(\cdot)$表示交叉熵，因此公式 (2) 的第一项和第二项分别表示分布内数据和异常值数据的 CCE。常数是联合训练的平衡权重。训练后，Softmax 分数 对于分布内数据会很高， 但对于，则会很低，从而得到对异常值的估计。尽管异常值数据集 与测试时使用的实际 OOD 数据不同，但深度神经网络分类器通过收紧决策边界以排除为 OE提供的异常值，学习到了区分 OOD 数据的能力。

2.3 能量分数

能量 分数是另一种用于分布外检测的度量，其定义为负的 e 自由能函数的值 自由能函数：

$$-E(x) = T\log\sum_{k=1}^{K}\exp\left(\frac{g_k(x)}{T}\right), \tag{3}$$

其中 是分类器对 并且 是温度参数 (2)。众所周知，Softmax 分数是能量分数 (2) 的特例，其中所有 logits 都加上了最大 logit 值的偏置。由于这种带偏置的评分函数不适用于分布外检测，因此据称能量分数比 Softmax 分数更具优势。

2.4 能量有界学习

虽然能量分数仅对已训练网络采用不同的评分方式，但能量有界学习（2）利用自由能函数进行训练。深度神经网络分类器经过训练后，会对分布内数据给出较高的能量分数，对异常值数据给出较低的分数。为此， 正则化损失 由以下公式中的两个平方铰链损失项定义：

$$\mathcal{L}_{energy} = \mathbb{E}_{(x_{in},y)\sim D_{in}}(\mathrm{Max}(0, E(x_{in}) - m_{in}))^2 + \mathbb{E}_{x_{OE}\sim D_{OE}}(\mathrm{Max}(0, m_{OE} - E(x_{OE})))^2. \tag{4}$$

其中，和分别是分布内数据和异常值数据的间隔超参数。统一分类损失与正则化损失的总损失可表示如下：

$$\mathcal{L}_{total} = \mathbb{E}_{(x_{in},y)\sim D_{in}}(\mathcal{H}(y, f(x_{in}))) + \beta \mathcal{L}_{energy}. \quad (5)$$

此处，是联合训练的平衡权重。在推理过程中，使用能量分数进行分布外检测。

3. 现场噪声暴露（ONE）

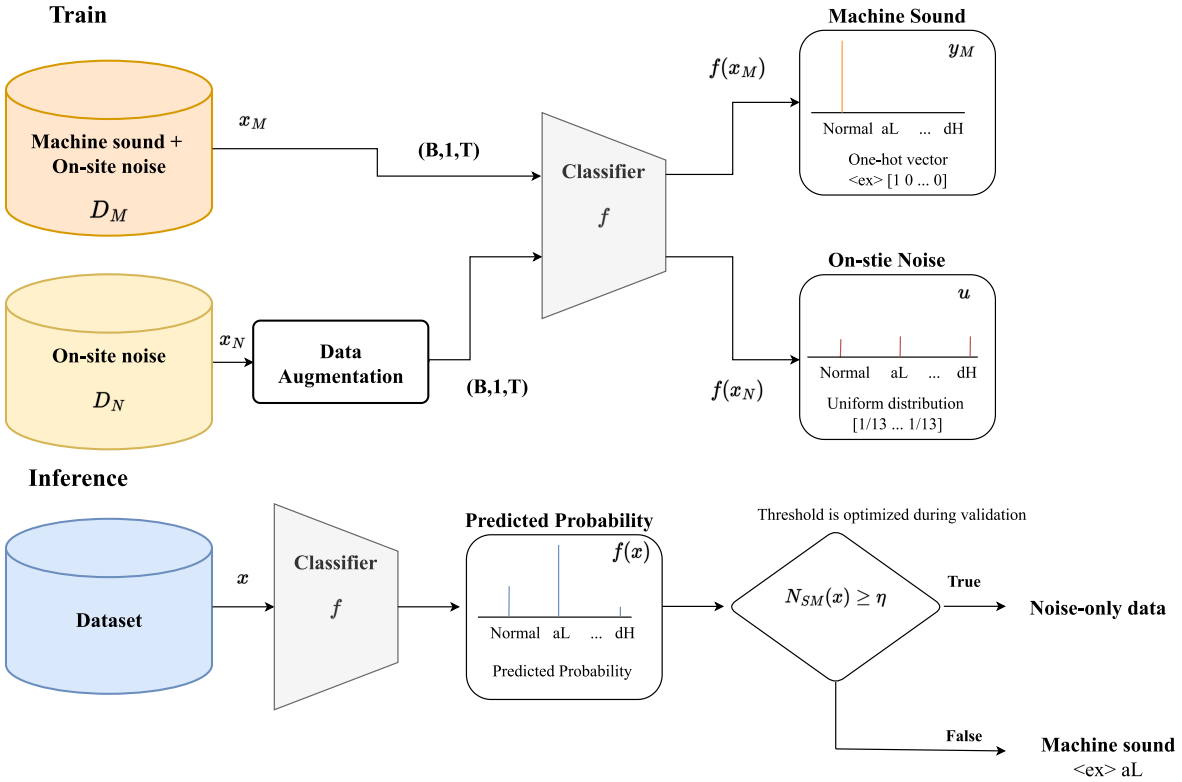

图 1 – ONE 的训练与推理流程

在工厂现场采集的数据可以分割为固定长度的信号片段。每个片段可能只包含背景噪声，我们希望将其判定为异常值。为提高检测性能，我们采用 OE 技术，使用机器停机期间测得的工厂背景噪声。

图 1 展示了 ONE 的训练和推理流程。用于训练深度神经网络模型的数据集由分布内数据 （包含受现场工厂噪声污染的机器声音）和纯噪声数据D_M（在纯噪声时段测得）组成。在训练阶段，从和中分别采样两个输入批次。两个批次的数据尺寸均为，其中表示批次大小，1 表示通道数（单声道），而 表示时间上的采样点数。在训练过程中，分类器根据输入 预测故障类型概率，其目标 由独热向量给出。与 OE 一样，训练分类器时要求其对纯噪声数据输出均匀分布。类似于公式 (2)，损失函数可定义为

$$\mathcal{L}_{NE} = \mathbb{E}_{(x_M,y_M)\sim D_M}(\mathcal{H}(y_M, f(x_M))) + \alpha \mathbb{E}_{x_N \sim D_N}(\mathcal{H}(u, f(x_N))). \quad (6)$$

我们将公式（7）的第二项记为噪声暴露损失。

At 在推理时，我们使用噪声分数 定义的纯噪声数据

$$N_{\mathrm{SM}}(x) = 1 - \mathrm{Max}(f(x)). \quad (7)$$

模型在推理时会计算每条数据的噪声分数，当噪声分数高于预定义阈值时，将其判定为纯噪声数据。当噪声分数低于阈值时，模型通过找到最大值对应的类别来对故障类型进行分类。该阈值通过验证集步骤确定，从而得到最优模型

performance（宏平均F1分数）可以获得，然后使用固定阈值进行测试 st.

一种可能的噪声暴露替代形式是使用自由能函数（FE）作为噪声分数。也就是说，

$$N_{FE}(x) = -E(x)\,. \tag{8}$$

公式（1）中提到的能量有界学习（EB）也可用于噪声暴露训练。在这种情况下，损失函数可定义为

$$\mathcal{L}_{EB} = \mathbb{E}_{(x_M, y)\sim D_M}(\mathcal{H}(y_M, f(x_M))) + \beta\mathcal{L}_{energy} \tag{9}$$

where $\mathcal{L}_{energy} = \mathbb{E}_{(x_M, y_M)\sim D_M}(\mathrm{Max}(0, E(x_M) - m_M))^2 + \mathbb{E}_{x_N\sim D_N}(\mathrm{Max}(0, m_N - E(x_N)))^2$,

对于间隔超参数 和 分别对应机器声音和纯噪声数据。对于能量有界学习，使用与公式 (8) 相同的噪声分数。

我们研究的噪声暴露的最后一种替代形式，是为一个模型提供训练集，使其将仅包含噪声的数据分类到一个单独的类别中（附加类别；AC）。针对这一目标的损失函数可写为

$$\mathcal{L}_{AC} = \mathbb{E}_{(x_M, y)\sim D_M}(\mathcal{H}(y_M, f(x_M))) + \mathbb{E}_{x_N\sim D_N}(\mathcal{H}(y_N, f(x_N))), \tag{10}$$

其中，目标概率向量 仅第（ ）类为 1，其余为 0。 $K+1$

4. 实验

4.1 数据集

表 1 – 玩具汽车数据集中各类别的数据数量

	正常	故障	噪声
训练集	200	600（每种故障类型各50个）	800
验证集	100	300（每种故障类型25个）	400
Test	100	300（每种故障类型25个）	400

表 2 – 玩具火车数据集中各类别的数据数量

	正常	故障	噪声
训练集	480	1440（每种故障类型120个）	1920
验证集	160	480（每种故障类型各 40 个）	640
Test	160	480（每种故障类型各40个）	640

在实验中，我们使用了ToyADMOS2 数据集 (5)，该数据集由来自两种玩具机器（玩具汽车和玩具火车）的声音以及环境噪声组成。在数据集中的五种不同机器型号类型和速度等级中，我们为玩具汽车和玩具火车数据均选择了型号 A 和速度 1。每台机器的状态标记为四种故障类型（a、b、c 和 d）以及三个损伤等级（低、中和高），因此包括正常状态在内共有 13 种状态。每条数据的时长为 12秒，采样率为16kHz。ToyADMOS2的噪声数据包括在四种不同环境中录制的工厂噪声信号：N1、N2、N3 和 N4。我们将这些称为噪声环境，以表示模型是通过相同还是不同的噪声数据进行训练和测试的。为了增加噪声数据的多样性，在训练步骤中，对每个工厂噪声信号应用了时间平移（0–2 秒）和音量扰动（0.5–2 倍）数据增强。

为模拟噪声环境，我们将机器声音与噪声数据以-10 dB至0 dB 的信噪比进行混合。考虑到现实情况下故障机器数据的稀缺性，我们使用了数量为故障数据四倍的正常数据。训练集、验证集和测试集数据集的详细信息见表 1 和表2。在本工作中，模拟纯噪声时段还需要纯噪声数据，因此我们通过收集 ToyADMOS2 中的噪声数据，单独准备了噪声数据集。

4.2 模型架构

表 3 – 模型结构

操作	输入	输出

梅尔频谱图	(,1,192000)B	(,1,128,374)B
批归一化	(,1,128,374)B	(,1,128,374)B
特征提取器	(,1,128,374)B	(,1024,1,1)B
压缩	(,1024,1,1)B	(,1024)B
线性分类器	(,1024)B	(,13)B
Softmax 激活	(,13)B	(,13)B

我们修改了 Koutini 提出的用于 ASC（声学场景分类）的残差网络 等人 （6）。模型结构以及输入和输出的大小如表3 所示。模型首先将原始音频信号转换为梅尔频谱图，然后通过批归一化进行标准化，并将其输入特征提取器。特征提取器的输出经过压缩后，被输入线性分类器层。最后，应用 softmax 激活函数以生成预测概率。由于共有 13 个机器状态类别，模型的输出形状为 ，批大小为。 (, 13)B E

4.3 训练集与验证集

我们对模型进行了 100 个训练轮次。机器声音和噪声的批次大小分别为 8，因此总批次大小为 16。我们使用 Adam 优化器，学习率从第 1 个训练轮次到第 30 个训练轮次设为 1e-4。到第 90 个训练轮次时，学习率线性降低至 1e-5，随后保持不变直至第 100 个训练轮次。

我们测试并比较了使用 ONE 训练的模型与仅使用 Softmax 分数、自由能量分数（FE）、能量有界学习（EB），以及将仅含噪声的时段归为附加类别（AC）的模型。所有技术的详细信息见表 4。

表 4——技术详情

技术	Loss	噪声分数	噪声暴露 data
Softmax 分数	CCE	负 MSP	不存在
噪声暴露（NE）	CCE、噪声暴露损失	负 MSP	存在
能量分数（FE）	CCE	负能量分数	不存在
能量有界学习（EB）	CCE，正则化损失	负能量分数	存在
含噪声数据的附加类别 (AC)	CCE	-	存在

为构建损失函数，平衡权重被设置为 用于噪声暴露，且用于能量有界学习。能量有界学习的间隔 和 被分别设置为和，分别 。温度参数 设为 1。每轮训练结束后都会对模型进行验证。噪声分数阈值 设定为优化模型性能，并在每轮训练中记录。经过 100 轮训练后，使用最佳模型参数和阈值进行测试。 α_0 α_1 m_M m_N -25 -7 T η

4.4 测试集

所有实验所使用的评估指标均为宏平均F1分数，即14个不同类别的F1分数的平均值（一个正常类别、一个噪声类别和12个故障类别）。由于噪声类别包含在分类标签中，因此可以通过单一指标同时评估故障类型分类和噪声检测性能。具体而言，宏平均F1分数定义为

$$F1_{Macro} = \frac{1}{14}\sum_{k=1}^{K} F1_k \ . \tag{11}$$

每个实验使用不同的随机种子重复进行三次，并展示分数分布及其平均值。

5. 结果

我们开展了三项不同的实验，以验证噪声暴露。在第 5.1 节中，我们使用来自同一环境的噪声训练和测试模型，将ONE与其他技术进行比较。在第 5.2 节中，我们考察训练集所使用的噪声与测试集所使用的噪声不同时的性能变化。最后，在第5.3节中，我们比较了针对现场噪声数据集和其他地点噪声数据集的噪声暴露性能。

5.1 在相同噪声环境下的性能

在第一个实验中，我们使用不同的现场噪声比较了五种不同的方法。来自同一环境的噪声被用于训练集和测试集，但两者之间没有重叠样本。图2展示了比较组的宏平均F1分数。我们可以看到，噪声暴露（NE）在玩具汽车上表现最佳。对于玩具火车，在噪声环境N1中，噪声暴露优于其他方法；但在噪声环境N2、N3和N4中，能量有界学习优于其他方法。尽管如此，噪声暴露在各项技术中仍表现出第二佳的性能。在玩具火车数据中也可以发现类似趋势（图3），其中噪声暴露和能量有界学习表现最佳。

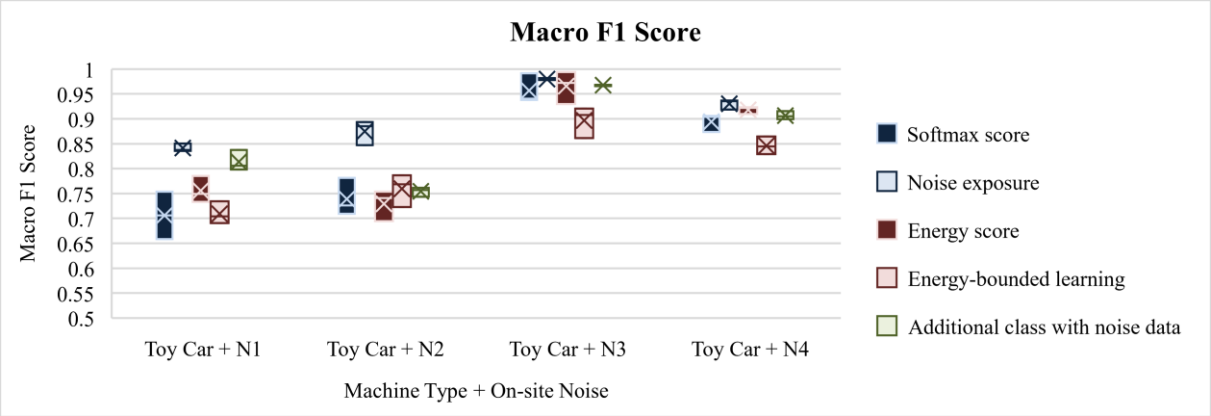

图 2 –不同模型在训练集和测试集所使用的噪声来自同一环境时的性能比较（玩具车数据）。（N1–N4：来自不同环境的噪声）

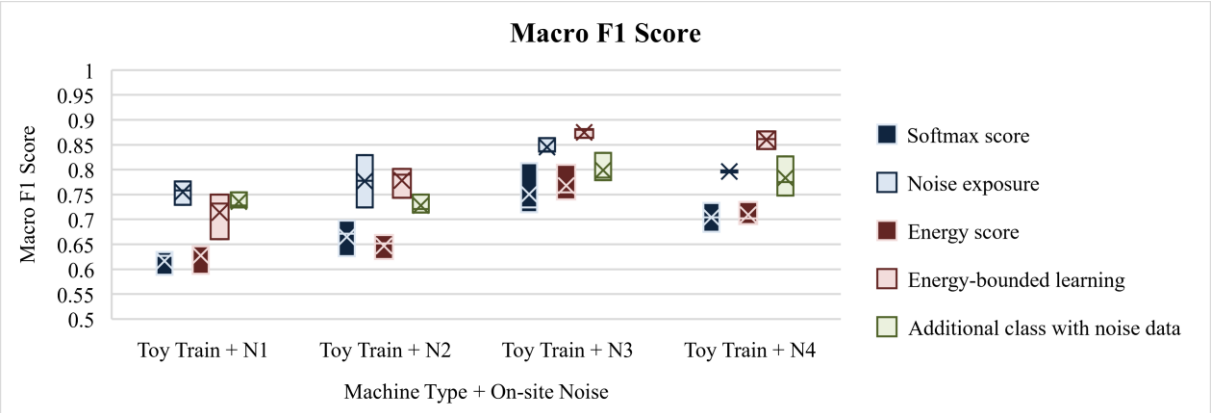

图 3 – 训练集和测试集使用来自相同环境的噪声时不同模型的性能比较（玩具火车数据）。（N1–N4：来自不同环境的噪声）

5.2 在未见噪声环境下的性能

当训练集和测试集使用来自不同环境的噪声时，测试结果如表 5和表 6所示。每行的第一个单元格分别表示训练集、验证集和测试集所使用的噪声环境。每一列代表所采用的技术。粗体文本突出显示每行的最佳情况。表5 和表 6显示，对于玩具汽车和玩具火车的大多数情况，噪声暴露的表现最佳。即使噪声暴露的分数不是最高，它仍然表现出第二好的结果。

性能。

T表 5 – 训练噪声来自不同环境时不同模型的性能比较 g
（验证集）和测试集来自不同环境（玩具车数据）。

训练噪声 / 验证噪声 / 测试噪声	Softmax 分数	噪声暴露	能量分数	能量受限学习	附加类别和噪声数据
N1 / N1 / N2	0.564	**0.777**	0.509	0.640	0.692
N1 / N1 / N3	0.490	**0.676**	0.463	0.594	0.617
N1 / N1 / N4	0.386	0.518	0.387	0.189	**0.521**
N2 / N2 / N1	0.536	**0.703**	0.543	0.487	0.427
N2 / N2 / N3	0.657	**0.799**	0.644	0.690	0.668
N2 / N2 / N4	0.343	**0.655**	0.322	0.218	0.377
N3 / N3 / N1	0.528	**0.601**	0.575	0.293	0.489
N3 / N3 / N2	0.534	**0.610**	0.553	0.336	0.611
N3 / N3 / N4	0.380	**0.397**	0.392	0.240	0.324
N4 / N4 / N1	0.483	**0.539**	0.512	0.387	0.348
N4 / N4 / N2	0.576	**0.590**	0.587	0.361	0.476
N4 / N4 /N3	**0.579**	0.577	0.551	0.377	0.522

表 6 – 训练集（验证集）和测试集使用的噪声来自不同环境时，不同模型的性能比较（玩具火车数据）。

训练集 噪声 / 验证噪声 / 测试噪声	Softmax 分数	噪声暴露 re	能量分数	能量有界学习	附加类别 with 噪声数据
N1 / N1 / N2	0.572	**0.779**	0.534	0.628	0.683
N1 / N1 / N3	0.589	**0.769**	0.591	0.587	0.687
N1 / N1 / N4	0.475	**0.734**	0.435	0.586	0.609
N2 / N2 / N1	0.493	**0.583**	0.483	0.415	0.443
N2 / N2 / N3	0.575	**0.694**	0.566	0.551	0.574
N2 / N2 / N4	0.441	**0.683**	0.442	0.541	0.551
N3 / N3 / N1	0.445	0.497	0.412	**0.521**	0.368
N3 / N3 / N2	0.418	**0.520**	0.407	0.519	0.414
N3 / N3 / N4	0.528	**0.720**	0.523	0.541	0.557
N4 / N4 / N1	0.367	**0.433**	0.379	0.311	0.262
N4 / N4 / N2	0.429	**0.537**	0.424	0.470	0.393
N4 / N4 /N3	0.420	**0.519**	0.429	0.367	0.389

5.3 使用现场噪声和其他地点噪声进行噪声暴露的性能

在本实验中，我们为训练集、验证集和测试集使用了不同的噪声数据集。表7和表8的第一列表示训练集和验证集阶段所使用的噪声，而其他列表示使用机器声音进行测试时的宏平均F1分数，这些机器声音混合了相应列第一个单元格所示的噪声。测试集纯噪声数据的噪声环境与测试集机器声音数据的噪声环境相同。因此，表7和表8的对角线（灰色）表示现场噪声暴露的结果，而非对角线表示其他地点噪声暴露的结果。使用玩具汽车和玩具火车数据集进行的测试表明，对于噪声暴露而言，现场噪声比其他地点的噪声更有益。在玩具汽车测试中存在一个例外情况（使用N3进行测试），但总体而言，暴露于相似类型噪声的模型对含噪数据更加鲁棒。这一结论看似显而易见，但也强调了我们无需收集各种噪声数据，只需将模型暴露于在工厂内测量的原位噪声中，就能稳健地训练模型。

表 7 – 通过噪声暴露训练的不同模型的性能比较
使用现场噪声（ONE）和其他地点噪声（玩具车数据）。

用于噪声暴露的噪声数据 噪声暴露	机器声音 使用 N1	机器声音 使用 N2	机器声音 使用 N3	机器声音 使用 N4
N1	**0.841**	0.865	0.985	0.925
N2	0.773	**0.875**	**0.988**	0.919

N3	0.701	0.778	0.980	0.889
N4	0.704	0.792	0.982	**0.930**

表 8 – 使用现场噪声（ONE）和其他现场噪声通过噪声暴露训练的不同模型的性能比较（玩具火车数据）。

用于的噪声数据 噪声暴露	机器声音 在 N1 条件下	机器声音 在 N2 下	机器声音 在 N3 条件下	机器声音 在 N4 条件下
N1	**0.755**	0.761	0.801	0.752
N2	0.732	**0.777**	0.776	0.778
N3	0.668	0.705	**0.845**	0.743
N4	0.625	0.714	0.743	**0.796**

6. 结论

本文将深度神经网络模型暴露于各种噪声中，以构建抗噪声分类器并检测纯噪声数据。针对这一目标，我们比较了五种不同的异常值暴露方法。通过噪声暴露，分类器在接受训练以对机器工况进行分类的同时，对纯噪声数据输出均匀的预测概率。异常值暴露方法的比较结果表明，无论机器声音和噪声的类型如何，噪声暴露的性能均为最佳或次佳。我们还使用不同的现场噪声和其他地点的噪声测试了噪声暴露。在现场噪声测试中，训练集和测试集使用了同一噪声数据集的不同划分；而在其他地点噪声测试中，则使用了不同的数据集。结果表明，使用现场噪声进行噪声暴露的效果优于使用其他地点噪声的测试结果。这些结果证明，在机器运行的同一环境中测得的噪声可用作噪声暴露数据集，并能提高故障分类器的鲁棒性。

致谢

本研究得到韩国政府（科学技术信息通信部）资助的韩国国家研究基金会（NRF）项目（编号：NRF-2020M2C9A1062710）支持，并得到韩国教育部资助、由韩国国家研究基金会（NRF）实施的BK21 Four计划支持。

参考文献

1. Hendrycks D、Gimpel K。神经网络中误分类样本和分布外样本检测的基线方法。载于：第 5 届国际学习表征会议（ICLR 2017），法国土伦，2017 年 4 月 24—26 日，会议论文集。2. Liu W、Wang X、Owens J、Li Y。基于能量的分布外检测。载于：神经信息处理系统进展。Curran Associates, Inc.；2020。第 21464–75 页。3. Hendrycks D、Mazeika M、Dietterich TG。通过异常值暴露进行深度异常检测。载于：第 7 届国际学习表征会议（ICLR 2019），美国路易斯安那州新奥尔良，2019 年 5 月。4. Primus P、Zwifl M、Widmer G。CP-JKU 在 DCASE’21 的投稿：通过代理异常值和基于语义对齐的域适应，改进用于机器状态监测的分布外检测器。DCASE2021 挑战赛。2021；5. Harada N、Niizumi D、Takeuchi D、Ohishi Y、Yasuda M、Saito S。ToyADMOS2：一种用于域偏移条件下异常声音检测的微型机器运行声音数据集。6. Koutini K、Eghbal-zadeh H、Dorfer M、Widmer G。感受野作为声学场景分类深度卷积神经网络中的正则化器。载于：2019 年第 27 届欧洲信号处理会议（EUSIPCO）。西班牙拉科鲁尼亚：IEEE；2019。第 1–5 页。