

基于随机配置网络的光纤入侵信号识别算法

盛智勇, 曾志强*, 曲洪权, 李伟

北方工业大学电子信息工程学院, 北京 100144

摘要 随机配置网络(SCN)通过引入不等式约束来限制输入权重和偏置的赋值,随着节点数量增加,网络能够逼近任意的数学函数和数据模型。在构建 SCN 的过程中,由于网络本身性质以及样本数据的不适定性和病态条件等问题会引起网络的过拟合,故提出一种基于 Dropout 技术的改进型 SCN 模型(Dropout-SCN)来自适应地约束输出权重分布和大小,以此来提高网络模型的识别精度。光纤数据验证的结果表明:与传统的 SCN 和 L2 范数正则化的 SCN 模型相比,Dropout-SCN 模型具有更低的测试误差,有效地减缓了网络过拟合问题,提高了对光纤预警系统(OFPS)中光纤入侵信号的识别准确率。

关键词 光通信; 随机配置网络; L2 正则化; Dropout 技术; 光纤预警系统; 信号处理

中图分类号 TP391

文献标识码 A

doi: 10.3788/LOP56.140602

Fiber Intrusion Signal Recognition Algorithm Based on Stochastic Configuration Network

Sheng Zhiyong, Zeng Zhiqiang*, Qu Hongquan, Li Wei

School of Electronic Information Engineering, North China University of Technology, Beijing 100144, China

Abstract A stochastic configuration network (SCN) introduces inequality constraints to limit the assignment of input weights and biases. The network can approximate arbitrary mathematical function and data model as the number of hidden nodes gradually increases. In the process of SCN construction, the properties of the network itself and the ill-posed and ill-conditioned problems of the sample data may cause over-fitting of the network model. This study proposes an improved SCN model based on the Dropout technology, called Dropout-SCN, to improve the recognition accuracy of the network model by adaptively constraining the output weight distribution. We then perform a verification using optical fiber data. Compared with the traditional SCN and L2 norm regularized SCN models, the Dropout-SCN model has a lower test error, which effectively prevents the network over-fitting problem and improves the recognition accuracy of the intrusion signals in the optical fiber pre-warning system.

Key words optical communications; stochastic configuration network; L2 regularization; Dropout technology; optical fiber pre-warning system; signal processing

OCIS codes 060.2310; 060.2370; 060.2380

1 引言

光纤预警系统(OFPS)可以实时监测运输管道状态,一旦管道发生泄漏或者是被人为破坏,光纤就将振动信号传回终端并准确地报告破坏位置^[1-3]。在光纤预警研究中,文献[4-5]研究了基于光纤信号的反射和干涉对光纤入侵信号进行直接提取的可行性,虽然在一定程度上可以检测到入侵信号,但是面

对背景噪声较强的低信噪比信号时并不能直接提取。文献[6]中提出了一种基于恒虚警率(CFAR)算法的光纤入侵信号检测方法,针对不同的背景噪声,该方法能够自适应地调节阈值,极大地提高了OFPS的检测性能。但在光纤信号识别方面,主要是通过信号进行特征提取,再经过分类器训练得到信号的识别模型^[7-9]。这类识别方法存在两个缺点:一是经过特征提取再进行分类识别耗时较多,二

收稿日期: 2018-12-12; 修回日期: 2019-01-21; 录用日期: 2019-02-25

基金项目: 国家自然科学基金(61571014,61601006)、北京自然科学基金(4172017)

* E-mail: 13101040127@mail.ncut.edu.cn

是有效的信号特征并不易提取,同时,固定的特征容易导致信号识别模型的适应性较差。因此,在OFPS中,一种适应性强、准确率高的快速识别方法对于光纤入侵信号的识别尤为重要。

神经网络中有许多非线性隐藏单元,这使得它们拥有非常强大的学习表示能力,可以逼近任何模型,并且能够学习到输入和输出之间的复杂关系^[10-11]。其中,随机算法由于其强大的快速学习能力以及较低的计算代价逐渐引起了人们的重视^[12]。Wang等^[13]提出了一种通过不等式约束随机配置输入权重和偏置,并且隐藏节点可变的随机配置网络(SCN)模型,有效地保障了随机算法模型的通用逼近性质,并降低了模型复杂度。其后续研究也进一步地证明了SCN建模的可行性^[14-15]。然而,在进行最后一层线性回归运算时,若特征矩阵存在秩亏,或奇异值之间相差很大时,SCN网络的稳健性以及泛化性急剧下降,引起网络的过拟合,使得网络模型测试性能不可靠。关于病态数据离散化问题的研究详见文献^[16-17]。为了计算得到一个更加稳定的解集,有必要对SCN网络线性回归部分进行正则化处理,以保证SCN模型的泛化能力。

Dropout是一种常见且有效的方法,用以阻止网络过拟合问题。该方法以一定的概率删除或保留节点,能够有效打破网络节点之间的共适性,有效提高网络的泛化能力^[18]。本文提出一种基于Dropout方法的改进SCN网络模型(Dropout-SCN)。在SCN网络隐藏层和输出层之间的线性

回归部分使用Dropout,利用一种近似等效替代的思想,以概率 p 对SCN网络隐藏节点进行Dropout,相当于对原SCN网络的输出权重作不同程度约束的L2正则化。相比较直接对SCN网络作L2正则化,Dropout-SCN对输出权重的约束更加严格,也更加有针对性。通过OFPS测试,Dropout-SCN相比较原SCN而言,能有效地抵制过拟合的出现,增强了网络的泛化能力,提高了OFPS的性能。

2 基于Dropout-SCN的光纤入侵信号识别模型

2.1 OFPS原理及应用

OFPS主要由分布式光纤传感器、数据采集处理模块、计算机(PC)三大部分组成。OFPS工作流程如图1所示,激光器产生两束高度相关的连续相干光,经过声光调制器(AOM)模块之后产生光脉冲,再经过掺铒光纤放大器(EDFA)和拉曼放大器(Raman)进入分布式光纤传感器,在未受到入侵源的振动时,光纤内部的光强分布相对平稳,当传感器受到外界入侵信号的影响时,两束相干光的相位发生变化。由光的干涉原理可知,相干光的相位改变会导致相应位置的光强发生变化。在接收端经过分束器(PBS)、光电转换器(APD)、模数转换(A/D)之后送入信号处理模块。在信号处理模块主要实现光纤入侵信号的检测和识别。在检测阶段,采用文献^[6]中提出的CFAR算法来实现对入侵信号的检测

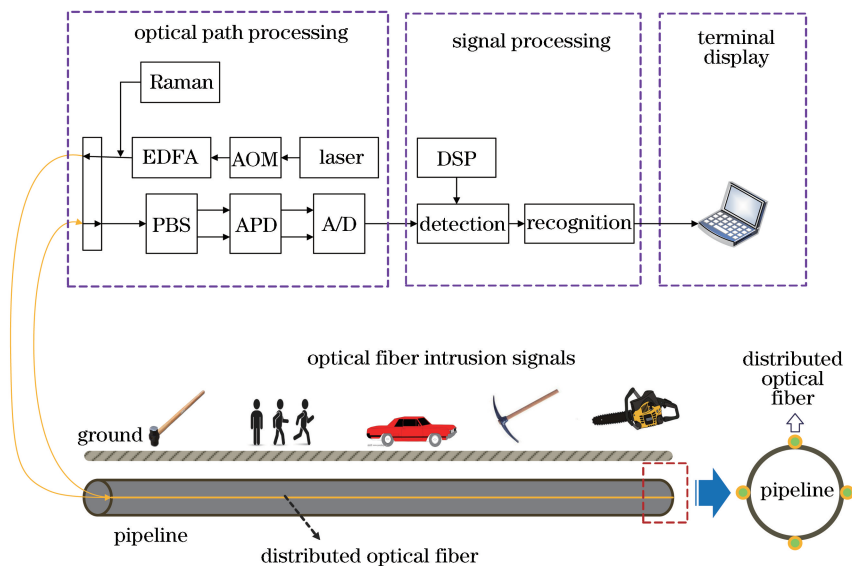


图1 光纤预警系统原理图

Fig. 1 Principle of optical fiber pre-warning system

和定位,即检测到前后时刻同一位置光强变化超过门限值时,便会向 PC 端提示报警位置。在识别阶段,通过对光纤振动信号进行数字信号处理(DSP),最终实现对入侵信号的识别。

OFPS 具有体积小、精度高、监测范围广、维护成本低等优点,被广泛应用于石油和天然气管道运输以及周界防护中。通过对分布式光纤的监测,实现对其周围环境的实时监控,在预防油气泄露、周界预警方面发挥了巨大的作用,保障了国家和人民财产的安全。经测试,OFPS 能够在长度 60 km、单根光纤周围 2 m 范围内有效地实现实时安全监测。由于入侵信号不一定是单一的纯振动信号,还常常伴随着多种环境噪声的叠加混合,仅仅通过“人工设计特征”方法来提取振动信号特征不能满足 OFPS 的需求。一方面,有效的人工设计特征不易提取,且效果不是很理想;另一方面,由于信号之间相互叠加、背景噪声复杂等问题,增加了对入侵信号的识别难度。在实际应用中,常常需要对多种入侵信号进行高精度的实时检测和识别,这向分类器的学习能力和逼近能力提出了挑战。一个拥有优秀的学习能力和较为简单网络结构的学习模型对 OFPS 的性能至关重要。

2.2 SCN 原理介绍

SCN 作为一种随机网络,广泛应用在处理数据的回归和分类问题上,并取得了很不错的效果。与传统的前向神经网络对比,SCN 在网络权重配置、网络节点数量、网络计算量等方面都有较大的优势。

2.2.1 SCN 训练过程

SCN 是一个标准的由输入层、隐藏层、输出层组成的 3 层前向反馈网络结构,其标准模型如图 2 所示,其中, g_j 为第 j 个隐藏节点的输出, L 表示隐藏节点的数量。在输入层采用一种具有监督学习规则的不等式约束随机赋值输入权重 w 和偏差 b ,具体的不等式赋值方式下一节会详细介绍。隐藏层采用 Sigmoid 作为激活函数,令 w_j 和 b_j 分别为第 j 个隐藏节点的输入权重和偏差,则第 j 个隐藏节点的输出为

$$g_j = \sigma \left(w_j^T \cdot \mathbf{X} + \begin{bmatrix} b_j \\ b_j \\ \vdots \\ b_j \end{bmatrix}_{N \times 1} \right), \quad (1)$$

式中: $\sigma(z) = \frac{1}{1 + \exp(-z)}$; “ \cdot ” 表示点积; \mathbf{X} 为样本输入。所以隐藏层输出矩阵 $\mathbf{h}_L(\mathbf{X})$ 为

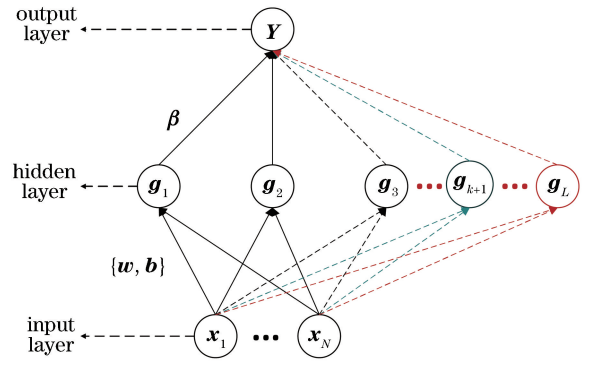


图 2 标准 SCN 模型

Fig. 2 Standard SCN model

$$\mathbf{H} = \mathbf{h}_L(\mathbf{X}) = [\sigma(w_1^T \mathbf{X} + b_1), \sigma(w_2^T \mathbf{X} + b_2), \dots, \sigma(w_L^T \mathbf{X} + b_L)]^T, \quad (2)$$

若隐藏层节点输出权重为 $\beta = [\beta_1, \beta_2, \dots, \beta_M]_{L \times M}$, M 为输出数量,则 N 个样本的隐藏节点输出为 $\mathbf{H} = [g_1, g_2, \dots, g_L]_{N \times L}$,整个网络的输出为

$$\mathbf{Y} = \mathbf{H} \cdot \beta. \quad (3)$$

在反向推导进行输出权值更新的时候,输出权重 $[\beta_1, \beta_2, \dots, \beta_{L-1}, \beta_L]$ 使用广义伪逆(Moore-Penrose generalized inverse)求得

$$\beta = \arg \min_{\beta} \|\mathbf{Y} - \mathbf{T}\|_{\text{F}}^2 = \arg \min_{\beta} \|\mathbf{H}\beta - \mathbf{T}\|_{\text{F}}^2 = \mathbf{H}^{\dagger} \mathbf{T}, \quad (4)$$

式中: \mathbf{T} 为标签矩阵; $\|\cdot\|_{\text{F}}$ 表示 Frobenius 范数; “ \dagger ” 表示伪逆。

得益于 SCN 独特的权值赋值方式,通过不等式约束随机配置输入权重 w 和偏差 b ,使用 Sigmoid 函数作为激活函数,并且输出层采用线性回归的思想,输出权重的计算使用最小二乘的线性方法,使得 SCN 在保障网络模型学习能力的同时还能够有效地降低网络模型训练时的计算复杂度以及缩短训练时间。SCN 最大的特点就是隐藏节点数量的可变性,在设置最大节点数量 L 、可容忍误差 ϵ 作为停止条件的前提下,网络节点数量随着训练的进行逐渐增加,直到满足任意一个停止条件时,网络停止训练,并输出最终模型,这极大地增加了网络节点数量的灵活性。

2.2.2 SCN 通用逼近性质

作为学习器的核心,通用逼近性质是保障训练过程中学习器性能的前提。对于 SCN 而言,在训练过程中,随着节点数量的增加,保障网络的误差趋于 0,无限逼近原函数,从而,得到一个学习能力强且拥有较强模型泛化能力的学习器是至关重要的。SCN 通用逼近性质由如下定理 1 进行约束:

$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} \left\| \mathbf{f} - \sum_j \boldsymbol{\beta}_j \mathbf{g}_j \right\|,$$

$\mathbf{e}_L = \mathbf{f} - \sum_{j=1}^L \boldsymbol{\beta}_j \mathbf{g}_j = [\mathbf{e}_{L,1}, \mathbf{e}_{L,2}, \dots, \mathbf{e}_{L,m}]$ 并且 $\mathbf{e}_0 = \mathbf{f}$, 其中: j 为隐藏节点; \mathbf{f} 为真实标签; \mathbf{e}_0 为初始残差; \mathbf{e}_L 为 L 个隐藏节点下的残差; m 为输出个数; $\mathbf{e}_{L,m}$ 为第 m 个输出的残差。

定理 1:

假设 $\text{span}(\Gamma)$ 在 L_2 空间中稠密, 并且 $\forall \mathbf{g} \in \Gamma$, $0 < \|\mathbf{g}\| \leq b_g$, 其中: $b_g \in \mathbf{R}^+$; Γ 为 L_2 的函数空间; b_g 是属于正实数空间的变量; $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$, $\mathbf{x}_i \in \mathbf{R}^d$. 令 $0 < r < 1$, 非负实数序列 $\{\mu_L\}$, 有 $\lim_{L \rightarrow \infty} \mu_L = 0, \mu_L \leq (1-r)$, 其中 r 为不等式约束系数。对于 $L=1, 2, 3, \dots$, 满足

$$\xi_{L,q} = \frac{[\mathbf{e}_{L-1,q}^T(\mathbf{X}) \cdot \mathbf{h}_L(\mathbf{X})]^2}{\mathbf{h}_L^T(\mathbf{X}) \cdot \mathbf{h}_L(\mathbf{X})} - (1-r-\mu_L) \mathbf{e}_{L-1,q}^T(\mathbf{X}) \mathbf{e}_{L-1,q}(\mathbf{X}) > 0, \quad (5)$$

式中: $q=1, 2, \dots, m$ 表示输出个数; $\mathbf{e}_{L-1,q}(\mathbf{X}) = [\mathbf{e}_{L-1,q}(\mathbf{x}_1), \mathbf{e}_{L-1,q}(\mathbf{x}_2), \dots, \mathbf{e}_{L-1,q}(\mathbf{x}_N)]$ 表示第 L 个节点增加前网络的残差; \mathbf{h}_L 是第 L 个节点的输出。需要注意的是, 随着 SCN 训练进程的继续, 隐藏节点数量逐渐增加, 模型的逼近能力也越来越强。与此同时, 不等式约束对于随机参数的约束也会越来越严格。具体而言, 对于任意新增加的隐藏节点 N_L , 其输入权重 \mathbf{w}_L 和偏差 b_L 的随机赋值空间会逐渐变大, 直到满足上述约束。对于任何一个新加入的节点权重和偏置, 只要满足(5)式, 即 $\xi_{L,q} > 0$, 随着节点数量的增加, SCN 可以逼近任意一个函数或者数据模型。

SCN 在面对大多数数据的时候都会有不错的表现。但由于其本身性质以及光纤数据的分布特性, 输入矩阵的数据会出现奇异值趋近于零的现象, 造成特征矩阵中的特征值相差很大。在进行最小二乘法求解输出权重的时候, 会出现病态条件的情况, 从而使得输出矩阵权值值域分布松散而不紧凑。SCN 计算中过分地拟合了训练数据中低贡献率的特征, 而忽略了具有共性的高贡献率的数据特征。这些往往会引起原网络的稳健性降低, 泛化能力下降。

2.3 Dropout-SCN 模型算法

Dropout 是一种常见的、有效的、用于抵制网络过拟合的技术。可以将 Dropout 理解成一种在隐藏节点中加入噪声的正则化方法^[18]。该方法将 Dropout 有效地应用于隐藏层中, 并通过一种类似于模型平均的思想重构整个网络的无噪声模。与传

统的正则化方法相比, Dropout 的不同之处在于节点选择的随机性。每一个节点的 Hidden 或 Visible 都是服从概率为 p 的伯努利分布: $R \sim \text{Bernoulli}(p)$, 由此保证每一次 Dropout 模型节点之间的随机性和独立性。

在随机配置网络中加入 Dropout 相当于在其隐藏层和输出层之间的线性回归部分删除和保留节点。Dropout-SCN 模型如图 3 所示。假设隐藏层中每一个节点保留和删除的概率都是固定的, 令每一个伯努利变量 $P(\delta_j=1)=p$, 即每一个节点被保留的概率为 p , 所以隐藏节点被删除的概率为 $P(\delta_j=0)=1-p=q$, 其中: δ_j 为伯努利变量; $\mathbf{f}_N = \sum_j \delta_j \boldsymbol{\beta}_j \mathbf{g}_j$ 为输出预测标签。所以 Dropout-SCN 网络的均方误差损失函数为

$$\min_{\boldsymbol{\beta}} E[\|\mathbf{f} - \mathbf{f}_N\|^2] = \min_{\boldsymbol{\beta}} E_{\delta_j \sim \text{Bernoulli}(p)} \left\| \mathbf{f} - \sum_{j=1}^L \delta_j \boldsymbol{\beta}_j \mathbf{g}_j \right\|^2, \quad (6)$$

式中: E 为误差期望; $\text{Bernoulli}(p)$ 为服从概率为 p 的伯努利分布; $E_{\delta_j \sim \text{Bernoulli}(p)}$ 为服从概率为 p 的伯努利变量的误差期望。

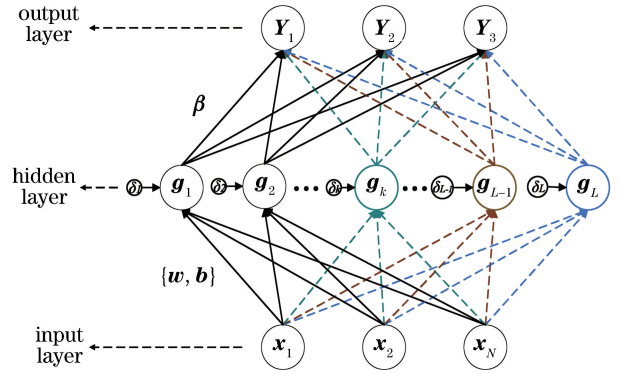


图 3 Dropout-SCN 模型: 展示了一个标准的多输入多输出网络模型

Fig. 3 Dropout-SCN model: standard multi-input and multi-output network model is presented

令 $\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_M]_{L \times M}$, $\mathbf{H} = [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_L]_{N \times L}$, L 为隐藏节点数, N 为样本数, M 是网络输出个数, Dropout 概率 p 等于一个固定值。可以进一步得到 Dropout-SCN 模型的目标函数为

$$\|\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_{L-1}, \boldsymbol{\beta}_L\| = \arg \min_{\boldsymbol{\beta}} \left\| \mathbf{f} - \mathbf{H} \hat{\boldsymbol{\beta}} \right\|^2 + \lambda \left\| \Gamma \hat{\boldsymbol{\beta}} \right\|^2, \quad (7)$$

式中: $\lambda = (1-p)/p$; $\boldsymbol{\beta} = p \hat{\boldsymbol{\beta}}$; $\Gamma = [\text{diag}(\mathbf{H}^T \mathbf{H})]^{1/2}$; $\text{diag}(\mathbf{H}^T \mathbf{H})$ 表示矩阵 $\mathbf{H}^T \mathbf{H}$ 的对角元素构成的对角矩阵。因此, 在 SCN 网络中使用 Dropout 方法的关

键是在线性回归部分进行节点的选择,与传统的岭回归相比较,Dropout-SCN 拥有一个特殊形式的 Γ 。Dropout-SCN 与标准 L2 正则化最大的不同在于通过计算不同节点之间的标准偏差 Γ 对输出权重 β 进行了不同程度的权值缩放。与 L2 正则化对所有权重进行相同的权值约束不同,Dropout-SCN 会对不同纬度的权重分别进行不同程度的约束。如果某一纬度上的数据与其他数据相比相差很大或者趋于奇异时,Dropout 算法会对其权重添加更多的限制。对该目标函数进行求解便可以得到 Dropout-SCN 网络新的输出权重 β 。这里直接使用标准最小二乘

法进行输出权重的求解。可以得到

$$\beta = [(\mathbf{H}^T \mathbf{H}) + \lambda \text{diag}(\mathbf{H}^T \mathbf{H})]^{-1} \mathbf{H}^T \mathbf{f}, \quad (8)$$

对角阵 $\mathbf{H}^T \mathbf{H}$ 比标准 L2 正则化对输出权重有更好的约束,这也是 Dropout-SCN 与标准 SCN、L2 正则化的 SCN 本质不同的地方,也是 Dropout-SCN 网络的核心问题所在。在此,给出 Dropout-SCN 网络模型:给定输入为 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$, $\mathbf{x}_i \in \mathbf{R}^d$, 输出为 $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$, $\mathbf{y}_i \in \mathbf{R}^m$; 隐藏最大节点数为 L_{\max} , 容忍误差为 ϵ , 最大搜索次数为 T_{\max} ; 随机权重范围 $\gamma = \{\lambda_{\min} : \Delta\lambda : \lambda_{\max}\}$ 。具体步骤如下:

1. Initialization settings $\mathbf{e}_0 = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]^T$, $0 < r < 1$, $\Omega, \mathbf{W} = []$;
2. **While** $L \leq L_{\max}$ and $\|\mathbf{e}_0\|_F > \epsilon$, **DO**
3. **For** $\lambda \in \gamma$, **DO**
4. **For** $k = 1, 2, \dots, T_{\max}$, **DO**
5. Randomly assign \mathbf{w}_L and b_L from $[-\lambda, \lambda]^d$ and $[-\lambda, \lambda]$;
6. Calculate $\xi_{L,q}, h_L$ according to Eqs. (2) and (5).
Let $\mu_L = (1-r)/(L+1)$;
7. **If**
8. $\min\{\xi_{L,1}, \xi_{L,2}, \dots, \xi_{L,m}\} \geq 0$
9. Save \mathbf{w}_L and b_L in \mathbf{W} , and $\xi_L = \sum_{q=1}^m \xi_{L,q}$ in Ω ;
10. **Else** go back to **Step 4**;
11. **End If**
12. **End For** (Corresponds to **Step 4**)
13. **If** \mathbf{W} is not empty
14. Find \mathbf{w}_L^* and b_L^* that maximize ξ_L in Ω ;
15. Set $\mathbf{H}_L = [\mathbf{h}_1^*, \mathbf{h}_2^*, \dots, \mathbf{h}_L^*]$;
16. **Break** (go to **Step 18**);
17. **Else** randomly take $\tau \in (0, 1-r)$, renew $r = r + \tau$;
18. Return to **Step 4**;
19. **End If**
20. **End For** (Corresponds to **Step 3**)
21. Calculate $\beta^* = [(\mathbf{H}^T \mathbf{H}) + \lambda \text{diag}(\mathbf{H}^T \mathbf{H})]^{-1} \mathbf{H}^T \mathbf{f}$ based on Eq. (8);
22. Compute $\mathbf{e}_L = \mathbf{e}_{L-1} - \mathbf{H}_L \beta^*$;
23. Renew $\mathbf{e}_0 = \mathbf{e}_L$; $L = L + 1$;
24. **End While**
25. **Return output** $[\beta_1^*, \beta_2^*, \dots, \beta_L^*], \mathbf{w}^* = [\mathbf{w}_1^*, \mathbf{w}_2^*, \dots, \mathbf{w}_L^*], b^* = [b_1^*, b_2^*, \dots, b_L^*]$;

其中 \mathbf{y}_i 为第 i 个样本标签, $\Delta\lambda$ 为步长, λ_{\max} 为最大随机参数取值, λ_{\min} 为最小随机参数取值, \mathbf{W} 为随机权值空间, Ω 为 ξ_L 赋值空间, L 表示隐藏节点数量, L_{\max} 为最大隐藏节点数量, d 为样本属性, τ 为随机

权值约束系数增量, $\|\mathbf{e}_0\|_F$ 表示对 \mathbf{e}_0 求 F 范数。

3 实验结果与分析

所有实验和仿真结果均来自 MATLAB-2017b,

Intel(R)Core(TM) i5-6200U CPU, Windows 10 系统平台。SCN 算法训练的停止条件为:最大节点数 $L_{\max}=100$ 和容忍误差 $\epsilon=0.01$ 。权值搜索参数设置为:最大搜索次数 $T_{\max}=100$,随机权重范围为 $\{0.5, 1, 5, 10, 30, 50, 100\}$, 不等式约束系数 $r=\{0.9, 0.99, 0.999, 0.9999, 0.99999, 0.999999\}$ 。

3.1 光纤数据预处理

在本实验中,使用了 3 种不同类型的光纤振动信号,分别为电钻产生的机械信号、铁锤敲打产生的敲击信号、镐刨挖地产生的挖掘信号,其归一化之后的原始信号如图 4 所示。在进行 Dropout-SCN 模

型网络训练之前,先对光纤原始信号进行预处理,以达到去直流,减少传感器的零点漂移现象,增大光纤信号信噪比的目的。对经过上述处理的数据进行快速傅里叶变换,将时域特征转换为频域特征,使得信号分布统一。同时选取傅里叶变换后的 128 维特征作为样本属性,并将其输入网络模型进行训练。实验使用了 1440 个光纤信号样本,其分配如下:1050 个训练样本、240 个测试样本、150 个验证集。其中每一类信号的比例均按照 1:1:1 的比例进行分配。例如:训练集 1050 个,即机械信号、敲击信号、挖掘信号各 350 个。

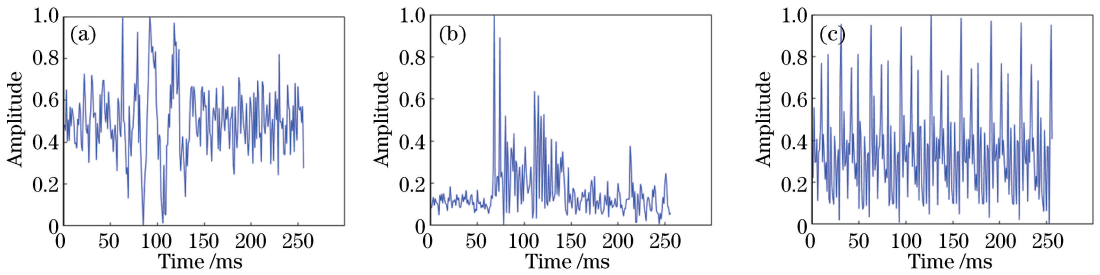


图 4 原始信号。(a)挖掘信号;(b)敲击信号;(c)机械信号

Fig. 4 Original signals. (a) Digging signal; (b) knocking signal; (c) electric drill signal

3.2 权值约束能力

实验比较了原始 SCN、Dropout-SCN 和 L2 正则化 SCN 3 种学习模型对输出权值的限制能力。在 30 个节点数且保证相同正则项系数的前提下,输出权值散点分布如图 5 所示。其中:纵坐标为权重大小,横坐标为第 L 个隐藏节点,空心圆圈代表原始 SCN 模型的权值分布散点图;从上到下第 1 条虚线为权值分布均值,其分布范围为 $0.1615 \sim 310.7753$,最大值与最小值之比为 1924;第 2 条虚线为 L2 正则化 SCN 网络输出权重均值,分布范围为 $0.0863 \sim 5.6430$,其中权重最大值与最小值之比为

65;第 3 条虚线为 Dropout-SCN 网络输出权重均值,分布范围为 $0.0248 \sim 0.9373$,权重最大值与最小值之比为 37。可以看到,相比较原始 SCN 模型和 L2 正则化 SCN 网络,Dropout-SCN 网络具有更加紧凑且更小的权值分布。这说明 Dropout-SCN 拥有更加严格的权值约束能力,这为网络的泛化能力提供了保障。

3.3 Dropout 概率对测试集影响

Dropout-SCN 网络拥有一个可以调节的超参数 p 。在本次实验中,为了探究不同的 Dropout 概率 p 对 Dropout-SCN 网络测试误差的影响,将网络节点数固定在 $L=70$,记录不同 Dropout 概率下的测试误差,实验结果如图 6 所示。在固定节点数量下,不同 Dropout 概率对网络测试误差的影响类似于一个下凹曲线,在中间拥有最低的测试误差($p=0.3 \sim 0.6$),而在两端测试误差较高,网络性能下降。由此可以看出:Dropout 概率为 $0.3 \sim 0.6$ 时能获得最优的测试误差。实际上,通常将 Dropout 概率设置为 $p=0.5$ 值。

3.4 测试误差实验

选用原始 SCN 以及 L2 正则化 SCN 模型作为对比进行实验,并使用固定的 Dropout 概率 $p=0.5$,即 $\lambda=(1-p)/p=1$ 。网络节点数量对测试误差的影响如图 7 所示。

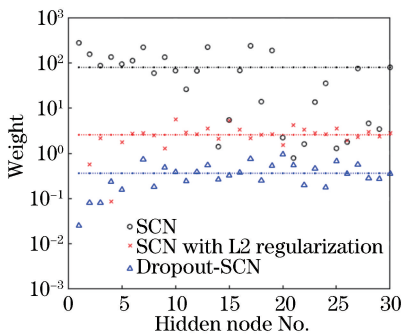


图 5 SCN、Dropout-SCN 和 L2 正则化 SCN 输出权重分布散点图

Fig. 5 Scatter plot of output weight distributions of SCN, Dropout-SCN and SCN with L2 regularization

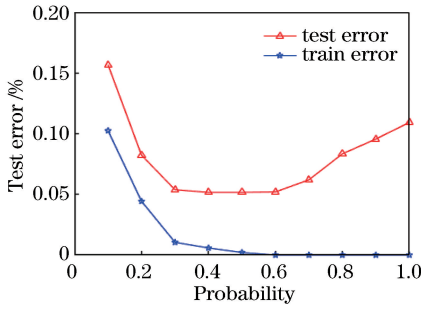


图 6 隐藏节点数 $L=70$ 条件下, Dropout 概率对测试误差的影响

Fig. 6 Effect of Dropout probability on test error under number of hidden nodes $L=70$

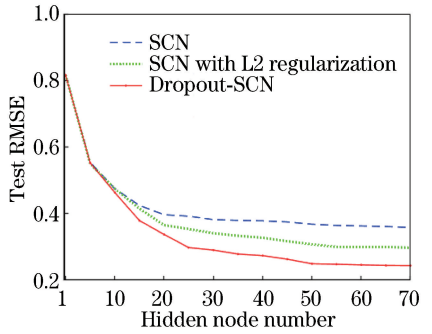


图 7 3 种网络模型的测试 RMSE

Fig. 7 Test RMSE for three network models

为了验证相比 SCN 网络, Dropout-SCN 的泛化能力确实有提升, 给出 3 种学习者模型在不同节点数 L 下的测试均方差 (RMSE)。随着节点数的增加, 3 种网络的测试均方差逐步下降, 与其他 2 种模型相比较, Dropout-SCN 拥有更低的测试误差, 这说明了模型对光纤入侵信号拟合得更好。另一方

面, 该结果也表明 Dropout-SCN 模型的稳健性比其他两种模型有所提高, 即 Dropout-SCN 网络的泛化性能更好, 能够有效地抵制网络过拟合现象。

3.5 对比实验

对 SCN、L2 正则化 SCN、Dropout-SCN 3 种网络在不同节点数 L , Dropout 概率为 $p=0.5$ 的条件下做了大量实验, 实验结果如表 1 所示。

表 1 光纤数据测试集结果

Table 1 Results of optical fiber test data

Network model	Hidden node number	Correct prediction number	Accuracy / %
SCN	50	214	89.17
SCN with L2 regularization	50	217	90.42
Dropout-SCN	50	222	92.50
SCN	70	217	90.42
SCN with L2 regularization	70	221	92.08
Dropout-SCN	70	227	94.58

表 1 展示了在节点数分别为 50 和 70 的情况下, 3 种网络模型对光纤信号的测试识别准确率。在节点数为 50 的情况下, SCN、L2 正则化 SCN、Dropout-SCN 3 种模型正确预测样本的数量分别为 214、217、222, 其准确率分别为 89.17%、90.42%、92.50%, 与其他 2 个网络相比较, Dropout-SCN 拥有更高的准确度。同样地, 在节点数 $L=70$ 的情况下, 各个网络的测试性能继续提高, 体现为正确预测样本数量的逐渐增加, 分别为 217、221、227, 准确率也进一步上升, 达到了 90.42%、92.08%、94.58%。可以看到, 在不同节点数量的情况下, Dropout-SCN 光纤信号识别模型都表现出了比其他 2 种网络模型更高的识别精度。这说明 Dropout-SCN 能够提高 OFPS 对入侵信号的识别性能, 并且有效地抵制了网络的过拟合, 增强了识别模型的泛化能力。

(Dropout-SCN), 用于减缓原始 SCN 存在的过拟合问题。通过理论证明: 在 SCN 中使用 Dropout 技术等效于对原始 SCN 作不同权值约束的 L2 正则化。Dropout-SCN 通过引入海森对角矩阵使得不同节点之间的标准差对输出权重大小进行了不同程度的权值约束。光纤数据实验结果表明: 与原始 SCN 模型相比, Dropout-SCN 能够降低光纤入侵信号约 4% 的测试误差, 有效地提升了光纤预警系统的泛化性能。

参 考 文 献

[1] Qu Z G, Zhou Y, Zeng Z M, *et al.* Detection of the abnormal events along the oil and gas pipeline and multi-scale chaotic character analysis of the detected signals [J]. Measurement Science and Technology, 2008, 19(2): 025301.

[2] Sun Q. Research on recognition method for Φ -OTDR optical fiber pre-warning pattern [D]. Tianjin:

4 结 论

提出了一种基于 Dropout 方法的 SCN 模型

- Tianjin University, 2015.
- 孙茜. Φ -OTDR 光纤预警系统模式识别方法研究 [D]. 天津: 天津大学, 2015.
- [3] Yang Y, Sha Z, Feng H, *et al.* Signal processing method of Φ -OTDR technology in early warning and leakage positioning of water pipelines [J]. *Laser & Optoelectronics Progress*, 2018, 55(4): 040607.
- 杨洋, 沙洲, 封皓, 等. 输水管道预警与泄漏定位中 Φ -OTDR 技术信号处理法 [J]. *激光与光电子学进展*, 2018, 55(4): 040607.
- [4] Wu Y H, Zhu S, Xu W, *et al.* Progress in distributed optical fiber crack sensing engineering [J]. *Laser & Optoelectronics Progress*, 2018, 55(9): 090002.
- 吴永红, 朱莎, 许蔚, 等. 分布式光纤裂缝传感工程应用研究进展 [J]. *激光与光电子学进展*, 2018, 55(9): 090002.
- [5] Li Q, Zhang C X, Li C S. Fiber-optic distributed sensor based on phase-sensitive OTDR and wavelet packet transform for multiple disturbances location [J]. *Optik*, 2014, 125(24): 7235-7238.
- [6] Qiu Z Z, Zheng T, Qu H Q, *et al.* A new detection method based on CFAR and DE for OFPS [J]. *Photonic Sensors*, 2016, 6(3): 261-267.
- [7] Yang D. Characteristics extraction and recognition research in optical fiber pre-warning signal [D]. Beijing: North China University of Technology, 2018.
- 杨丹. 光纤振动信号特征提取与识别算法研究 [D]. 北京: 北方工业大学, 2018.
- [8] Wang Z Y, Pan Z Q, Ye Q, *et al.* Fast pattern recognition based on frequency spectrum analysis used for intrusion alarming in optical fiber fence [J]. *Chinese Journal of Lasers*, 2015, 42(4): 0405010.
- 王照勇, 潘政清, 叶青, 等. 用于光纤围栏入侵告警的频谱分析快速模式识别 [J]. *中国激光*, 2015, 42(4): 0405010.
- [9] Wang Y H, Liu J W, Zhang M Y, *et al.* Test on axial force of double-walled mode pipe pile based on fiber Bragg grating [J]. *Laser & Optoelectronics Progress*, 2018, 55(11): 110605.
- 王永洪, 刘俊伟, 张明义, 等. 基于光纤布拉格光栅的双壁模型管桩轴力测试 [J]. *激光与光电子学进展*, 2018, 55(11): 110605.
- [10] Guo X L. Study on interpolation and approximation problem of neural networks [D]. Yinchuan: Ningxia University, 2017.
- 郭小乐. 关于神经网络插值与逼近问题的研究 [D]. 银川: 宁夏大学, 2017.
- [11] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors [J]. *Nature*, 1986, 323(6088): 533-536.
- [12] Mahoney M W. Randomized algorithms for matrices and data [J]. *Foundations and Trends[®] in Machine Learning*, 2011, 3(2): 647-672.
- [13] Wang D H, Li M. Stochastic configuration networks: fundamentals and algorithms [J]. *IEEE Transactions on Cybernetics*, 2017, 47(10): 3466-3479.
- [14] Wang D H, Cui C H. Stochastic configuration networks ensemble with heterogeneous features for large-scale data analytics [J]. *Information Sciences*, 2017, 417: 55-71.
- [15] Wang D H, Li M. Robust stochastic configuration networks with kernel density estimation for uncertain data regression [J]. *Information Sciences*, 2017, 412/413: 210-222.
- [16] Li H. Singular values correction methods with double parameters for ill conditioned problems and their applications [D]. Zhengzhou: PLA Strategic Support Force Information Engineering University, 2018.
- 李豪. 病态问题的奇异值双参数修正解法及其应用 [D]. 郑州: 战略支援部队信息工程大学, 2018.
- [17] Liu T, Dai Y K, Yang Y X, *et al.* Dynamic magnetoencephalography source reconstruction based on singular value decomposition and minimum L_2 -norm estimation [J]. *Chinese Journal of Stereology and Image Analysis*, 2015, 20(3): 235-241.
- 刘婷, 戴亚康, 杨莹雪, 等. 基于奇异值分解和最小 L_2 范数解的动态脑磁源重建 [J]. *中国体视学与图像分析*, 2015, 20(3): 235-241.
- [18] Sutskever I, Hinton G, Krizhevsky A, *et al.* Dropout: a simple way to prevent neural networks from overfitting [J]. *Journal of Machine Learning Research*, 2014, 15(1): 1929-1958.