

# 基于高斯神经元的自组织映射网络研究

刘 达, 陈松灿

(南京航空航天大学计算机科学与技术学院, 南京 211106)

**摘 要:** 自组织映射网络(Self-organizing map network, SOM)是一种经典的无监督学习方法,具有自组织和联机学习功能。由于其简明与实用等特点,不断涌现出SOM变体以适应各类问题。然而,这些工作基本都采纳了确定性神经元建立网络,忽略了数据本身隐含的不确定性信息,导致这些模型的结果缺乏由置信度反映的可解释性,意味着SOM神经元的不确定性刻画能力不足。本文提出了一种高斯神经元SOM网络(Gaussian neuron som network, GNSOM),其神经元节点不再是确定性的,而是建模为高斯分布的高斯神经元,为SOM配备了不确定性功能用于表述数据的不确定性。在实现时,将输入数据同样高斯化,并用Jensen-Shannon(JS)散度代替SOM学习中的欧氏距离作为GNSOM学习中的相似性匹配度量,由此获得了不确定性表示。实验结果表明,GNSOM具有更好的训练效果,并能通过神经元节点的协方差矩阵反映数据的不确定性。由于这种对神经元的高斯化独立于SOM本身,因此能拓展应用于其他神经元模型。

**关键词:** 无监督学习;自组织映射网络;数据不确定性;高斯神经元自组织映射;JS散度

**中图分类号:** TP301.6      **文献标志码:** A

## Research on Self-organizing Map Network Based on Gaussian Neuron

LIU Da, CHEN Songcan

(College of Computer Science and Technology, Nanjing University of Aeronautics & Astronautics, Nanjing 211106, China)

**Abstract:** Self-organizing map network (SOM) is a classic unsupervised learning method with self-organizing and online learning functions. Due to its simplicity and practicality, SOM variants have been emerging to adapt to various problems. However, these work basically adopts deterministic neurons to build networks, ignoring the uncertainty information implicit in the data itself. This results in a lack of interpretability reflected by confidence in the results of these models, implying that the uncertainty characterization ability of SOM neurons is insufficient. This article proposes a new variant of SOM, called the Gaussian neuron SOM network (GNSOM). Its neuron nodes are no longer deterministic, but modeled as Gaussian neurons with Gaussian distribution. Thus, SOM is equipped with an uncertainty function to express the uncertainty of the data. In implementation, the input data are also Gaussianized, and the Jensen-Shannon (JS) divergence is used to replace the Euclidean distance as the similarity matching metric in GNSOM learning, thereby obtaining the uncertainty representation. The experimental results show that GNSOM has a better training effect, and can reflect the uncertainty of the data through the covariance

matrix of the neuron node. Since this Gaussization of neurons is independent of SOM itself, it can be extended to other neuron models.

**Key words:** unsupervised learning; self-organizing map network; data uncertainty; Gaussian neuron self-organizing map; JS divergence

## 引言

自组织映射网络(Self-organizing map network, SOM)是一种通过无监督竞争学习方式将数据映射到规则二维网格以获得高维数据的低维离散映射的数据分析方法。SOM<sup>[1]</sup>是由芬兰神经网络专家Kohonen所提出,故又被称为Kohonen网络。作为一大类竞争学习网络的代表,SOM自提出以来因其算法实现原理简单、直觉,同时具有可视化和联机学习等众多功能,广泛地受到研究者持续关注 and 跟进,并已应用于金融、医疗和工业等众多领域<sup>[2-4]</sup>。为进一步适应多样化的现实问题,不断涌现出了相应的SOM变体,例如针对层次型数据提出的层次自组织映射网络<sup>[5-6]</sup>以及为了网络结构更灵活提出的自组织增量神经网络<sup>[7-8]</sup>等。至今,仍有大量与SOM相关的工作被发展和开拓,典型的如将其用于动态环境的连续学习<sup>[9]</sup>、异常检测<sup>[10]</sup>或深度化学习<sup>[11]</sup>等。通常现实获取的数据普遍存在不确定性,例如因测量误差引起的不确定性;图像获取过程中设备内在因素带来的噪声影响;问卷类方法收集的数据受来源主观影响产生的不确定性等。故在机器学习领域的发展过程中,关于数据不确定性的研究已成为其中不可或缺的组成部分。在数据带有不确定性的情况下,学习过程中对所有特征一视同仁并不合理,不确定性小的特征应该更受关注。然而在今日诸多SOM相关工作中,基本都采纳了确定性神经元建立网络,忽略了数据本身隐含的不确定性信息<sup>[12-16]</sup>。这意味着如下的隐含假设:输入数据是确定性的,显然对于大部分实际情况,未必成立。本文提出了针对性的高斯神经元SOM网络(Gaussian neuron self-organizing map network, GNSOM),为神经元赋以方差项刻画的数据不确定性,通过学习数据隐含的不确定性信息来提高学习过程的可靠性与合理性,而新增的方差项可加深对神经元的直觉认知,提高网络对输出结果的可解释性。

## 1 自组织映射网络

SOM网络模型包括两层:一是输入层,对应输入数据;二是竞争层,即神经网络。竞争层由多个带有权向量的神经元组成,且神经元被放置在网格的节点上。网络通常采用一维或二维设置,少数情况下会出现三维或更多维设置。图1(a)为一维网格模型,图1(b)为最常见的二维网格模型。SOM网络的一个显著特点是邻域学习,其竞争层神经元各有自己的邻域,竞争结果是获胜神经元及其邻域中的节点进行权值更新,亦即在参数更新步骤中,既更新获胜神经元的权值,又更新邻域内其余节点的权值。直觉而言,邻域内所有神经元都受到了不同程度的刺激。

现假设  $X = \{x_1, \dots, x_n\} \in \mathbb{R}^{D \times n}$  是一组给定的训练样本,  $W = \{w_1, w_2, \dots, w_M\} \in \mathbb{R}^{D \times M}$  对应SOM神经元权向量,不失一般性,设  $M = m^2$ ,  $m$  为网格边长,则SOM执行流程如下:

### 算法1 自组织映射网络

输入:数据集  $X$ , 网格边长  $Size$ , 特征数  $f$ , 邻

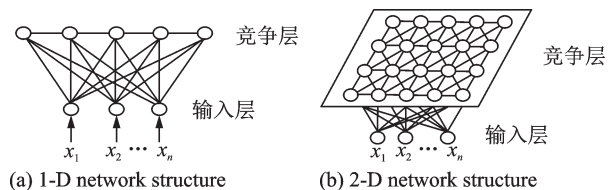


图1 SOM常见网络结构

Fig.1 Common network structures of SOM

域初始值  $\sigma_0$ , 学习率初始值  $\alpha_0$ , 迭代次数  $T_{\max}$

输出: 训练好的神经网络

(1) 参数初始化。随机选择输入数据对权向量进行初始化, 并指定初始邻域半径  $\sigma(0)$ , 初始学习率  $\alpha(0)$ , 网格边长  $m$  以及最大迭代次数  $T_{\max}$ 。

(2) for epoch = 1 to  $T_{\max}$

① 从数据集中随机选取数据  $x(i)$

② 利用式(1)计算输入数据分布与分布神经元的相似度, 并确定获胜神经元

$$i(x) = \operatorname{argmin}_j \|x(i) - \mathbf{w}_j\|^2 \quad j = 1, 2, \dots, M \quad (1)$$

③ 利用式(2)对获胜神经元及其邻域内神经元的参数进行更新, 其中  $h_{j,i(x)}(n)$  是获胜神经元  $i(x)$  的邻域函数,  $d_{j,i}$  是获胜神经元与邻域内神经元在网格中的距离

$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + \alpha(n) h_{j,i(x)}(n) (\mathbf{x} - \mathbf{w}_j(n)) \quad (2)$$

$$h_{j,i(x)} = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2}\right) \quad (3)$$

④ 利用式(4,5)对学习率及邻域范围进行动态调整

$$\alpha(n) = \alpha(0) \exp\left(-\frac{n}{T_{\max}}\right) \quad (4)$$

$$\sigma(n) = \sigma(0) \exp\left(-\frac{n}{T_{\max}}\right) \quad (5)$$

(3) end for

## 2 高斯神经元自组织映射网络

从 SOM 的更新(2)可发现神经元权向量实质是权向量朝数据向量移动, 实现对数据的拟合和数据分布的表达<sup>[4]</sup>。此外, 权向量的更新可看作是  $\|\mathbf{x} - \mathbf{w}\|^2$  作为优化目标联机最小化的更新 ( $\mathbf{w}$  是选定神经元的权向量)<sup>[17]</sup>。

由 SOM 的算法流程可见其对数据的表述是确定性过程, 无法直接反映数据隐含的不确定性, 并且其诸多变体也缺乏此刻画能力, 本文旨在弥补这一不足, 提出 GNSOM 来解决这一问题。存在诸多刻画不确定性的方式<sup>[18-21]</sup>, 而高斯分布因其简单有效是其中最常用的方式, 为此采用高斯分布来创建神经元, 并以方差来反映数据的不确定性。具体实现时, 将输入数据同样高斯化处理, 将单点数据转换成分布, 如此使原有 SOM 所用的欧氏距离或余弦距离不再适合作为相似性度量, 因此新的 GNSOM 中的相似性替换成分布间散度实现对多个输入分布的拟合。存在多种分布间度量, 常用的如 Kullback-Leibler (KL) 散度和 Jensen-Shannon (JS) 散度<sup>[22]</sup>。

现假设  $X$  是一组给定的样本, 每一数据拥有对应的均值向量  $\mu_i$  和方差向量  $\sigma_i, i = 1, 2, \dots, n$ 。同时 GNSOM 网络神经元拥有对应的均值向量  $\mu'_j$  和方差向量  $\sigma'_j, j = 1, 2, \dots, M$ 。

2 个分布  $Q$  与  $P$  之间 KL 散度定义为

$$D_{\text{KL}}(P||Q) = \int P(x) \lg \frac{P(x)}{Q(x)} dx$$

当 2 个分布都是高斯分布, 即  $P \sim N(\mu, \Sigma)$  以及  $Q \sim N(\mu', \Sigma')$ , KL 散度简化为

$$D_{\text{KL}}(P||Q) = \frac{1}{2} \left\{ \lg \frac{|\boldsymbol{\Sigma}'|}{|\boldsymbol{\Sigma}|} - n + \text{tr}(\boldsymbol{\Sigma}'^{-1} \boldsymbol{\Sigma}) + (\boldsymbol{\mu}' - \boldsymbol{\mu})^T \boldsymbol{\Sigma}'^{-1} (\boldsymbol{\mu}' - \boldsymbol{\mu}) \right\}$$

在GNSOM的每次更新中,利用梯度下降最小化分布间的KL散度得梯度为

$$\begin{aligned} \frac{\partial D_{\text{KL}}(P||Q)}{\partial \boldsymbol{\Sigma}'} &= \frac{1}{2} \left\{ \boldsymbol{\Sigma}'^{-1} - \boldsymbol{\Sigma} \boldsymbol{\Sigma}'^{-1} \boldsymbol{\Sigma}'^{-1} - \boldsymbol{A} \boldsymbol{A}^T \right\} \\ \frac{\partial D_{\text{KL}}(P||Q)}{\partial \boldsymbol{\mu}'} &= \frac{1}{2} \left\{ \left( \boldsymbol{\Sigma}'^{-1} + (\boldsymbol{\Sigma}'^{-1})^T \right) (\boldsymbol{\mu}' - \boldsymbol{\mu}) \right\} \end{aligned}$$

式中  $\boldsymbol{A} = \boldsymbol{\Sigma}'^{-1} (\boldsymbol{\mu}' - \boldsymbol{\mu})$ 。然而实现时,发现更新方差参数时,分母中方差的高次项会使更新过程出现不稳定的情况。另因KL散度具有不对称性,并不适合作为相似性度量。为此采用KL散度对称化后的JS散度,其定义为

$$\text{JS}(P||Q) = \frac{1}{2} \text{KL} \left( P \parallel \frac{P+Q}{2} \right) + \frac{1}{2} \text{KL} \left( Q \parallel \frac{P+Q}{2} \right) \quad (6)$$

当  $\boldsymbol{\Sigma} = \text{diag} \{ \sigma_1^2, \sigma_2^2, \dots, \sigma_n^2 \}$ , 即协方差矩阵为对角阵时,式(6)简化为

$$\text{JS}(P||Q) = \frac{1}{2} \sum_{i=1}^n \left( \lg \frac{\sigma_i'^2 + \sigma_i^2}{4\sigma_i'\sigma_i} + 2 + \frac{4(\mu_i' - \mu_i)^2}{\sigma_i'^2 + \sigma_i^2} \right) \quad (7)$$

式(7)中最后一项是加权项,会将较大的权重分配给不确定性较小的维度,这符合直觉。相关梯度计算如下

$$\frac{\partial \text{JS}(P||Q)}{\partial \sigma_i'} = \frac{\sigma_i'^4 - \sigma_i^4 - 8(\mu_i' - \mu_i)^2 \sigma_i'^2}{2(\sigma_i'^2 + \sigma_i^2)^2 \sigma_i'} \quad (8)$$

$$\frac{\partial \text{JS}(P||Q)}{\partial \mu_i'} = \frac{4(\mu_i' - \mu_i)}{\sigma_i'^2 + \sigma_i^2} \quad (9)$$

值得注意的是,方差项梯度受方差向量和均值向量共同影响,说明GNSOM可学习到数据中隐含的不确定性信息。根据梯度下降原则,结合式(2)的更新形式,易得参数迭代更新公式为

$$\mu_i'(n+1) = \mu_i'(n) - \alpha(n) h_{j,i(x)}(n) \frac{\partial \text{JS}(P||Q)}{\partial \mu_i'} \quad (10)$$

$$\sigma_i'(n+1) = \sigma_i'(n) - \alpha(n) h_{j,i(x)}(n) \frac{\partial \text{JS}(P||Q)}{\partial \sigma_i'} \quad (11)$$

算法2给出了GNSOM的具体学习流程。

#### 算法2 高斯神经元自组织映射网络

输入:数据集  $\boldsymbol{X}$ , 网格边长 Size, 特征数  $f$ , 邻域初始值  $\sigma_0$ , 学习率初始值  $\alpha_0$ , 迭代次数  $T_{\max}$

输出:训练好的分布神经元

(1)各项参数初始化

(2)for epoch = 1 to  $T_{\max}$

①从数据集中随机选取数据  $\boldsymbol{x}_i$ ;

②利用式(7)计算输入数据分布与分布神经元的相似度,并确定获胜神经元;

③利用式(10-11)对获胜神经元及其邻域内神经元的参数进行更新;

④对学习率及邻域范围进行调整。

(3)end for

### 3 实验与结果分析

本节通过实验验证上文提出的 GNSOM 的有效性。实验中仅与 SOM 作比较,理由是首先对数据点和权向量的高斯化处理是对神经网络基本单元的改造和拓展,独立于其他 SOM 变异,具有一般性,因此能直接集成到任何对 SOM 的改进变异。其次 GNSOM 对 SOM 性能的提升直觉上同样能提升其他变异模型的行为,这归咎于高斯化与现有 SOM 改进策略间的正交互补性。

#### 3.1 实验设置

数据集包括两类,一类是向量型数据集,如 iris、seeds、wine 和 pendigits;另一类是带有不确定性的分布型数据集。因为缺少现成的分布型数据集,实验中利用变分自编码器(Variational autoencoder, VAE)将 MNIST 和 FASHION-MNIST 数据集建模成带有不确定性信息的分布数据,数据的维度是 16 维。SOM 和 GNSOM 属于无监督学习方法,但在训练结束后可利用标签给每个神经元节点赋以标签,做法是将神经元的标签设置为投影到该神经元输入数量最多的类别,然后计算标签预测的准确率。

#### 3.2 结果分析

对向量型数据集,GNSOM 采用两种处理方式:一种仅考虑确定性数据,即训练过程中协方差矩阵不参与更新,这种情况下 GNSOM 即为 SOM,所以省略结果展示;另一种处理方式是给这些数据加上小方差扰动分布化再训练,结果如表 1 所示,数据集的分类精度都有提高。这是因为在训练过程中,方差会逐渐学习到关于数据特征的重要性信息。当投影到某个神经元的数字中有某些特征的值非常接近时,直觉上就可获知这些特征对该神经元的重要程度更高,反映到方差上就是这些特征具有较小的方差,而式(7)表明 GNSOM 中方差小的特征在确定获胜神经元时更重要。也即 SOM 对每个特征的重视程度无差异,而 GNSOM 对不同特征的重视程度会在训练过程中产生变化,符合直觉。

表 1 实验结果

数据集	iris	wine	seeds	pendigits	MNIST	FASHION_MNIST	%
SOM	93.20±1.40	91.08±1.37	87.42±1.43	92.30±0.14	88.24±0.47	75.84±0.48	
GNSOM	97.58±0.71	94.88±1.47	90.71±1.55	92.82±0.58	90.63±0.64	76.97±0.34	

下面分析由 VAE 网络生成的数据分布的实验结果。从表 1 可见准确率有明显提高,这直接说明了 GNSOM 确实能利用数据的不确定性信息。为了验证该方法在噪声干扰下的效果,给两种图片数据集分别加上了均值为 0.1、方差为 0.01 的高斯噪声(以-G 结尾)以及概率为 0.03 的椒盐噪声(以-SP 结尾),实验结果如表 2 所示,可见 GNSOM 受噪声的影响更小。尝试将均值向量和方差结合作为 SOM 输入的实验,最后分类准确率分别为 88.76%±0.27% 和 76.42%±0.21%。此外只用类别‘0’训练模型,其他类别当作异常,用训练集平均量化误差乘以比率作为阈值,进行异常检测。二者分别选择最佳的比率范围,得到精度最大值范围如图 2 所示,GNSOM 明显优于 SOM。在这种设置下,也将均值向量和方差结合进行了实验,SOM 精度仅有 50% 左右,相当于随机判断。从这些结果可看出,GNSOM 对不确定性的学习能力明显胜过 SOM,而在部分情况下,不恰当地采用不确定性会对 SOM 行为产生负面影响。

表 2 噪声干扰实验结果

数据集	MNIST-G	MNIST-SP	FASHION_MNIST-G	FASHION_MNIST-SP	%
SOM	86.36±0.28	86.96±0.37	70.41±0.44	70.07±0.43	
GNSOM	89.52±0.18	89.71±0.39	74.71±0.30	74.94±0.13	

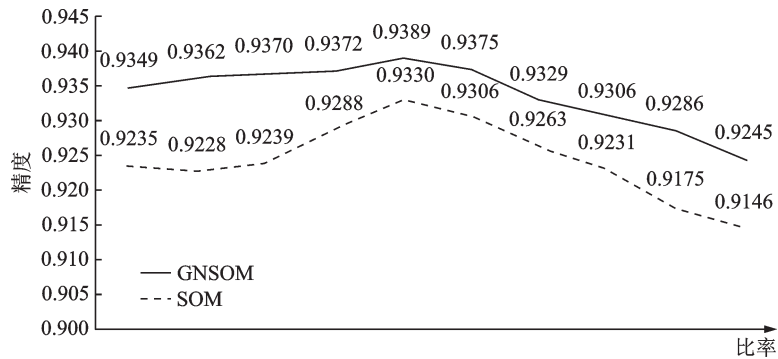


图2 两种方法在MNIST数据集上异常检测精度

Fig.2 Anomaly detection accuracy of two methods on MNIST dataset

### 3.3 不确定性分析

(1)不确定性的意义。分别从用MNIST和FASHION\_MNIST数据集训练好的GNSOM网络模型中选取两个节点,并根据方差大小确定这4个节点不确定性最高和最低的特征,然后给这2个特征附加同等程度的扰动。如图3所示,图3(a,b)左边对应节点原权向量重建的图片;中间对应节点不确定性最大的特征经扰动后重建的图片;右边对应节点不确定性最小的特征经扰动后重建的图片。可见经过相同程度

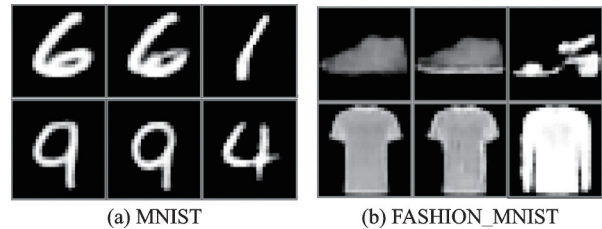


图3 GNSOM的权向量重构图

Fig.3 Image of GNSOM's weight vector reconstruction

的扰动,中间的重建图片变化较小仍维持原来的类别,而右边的重建图片变化极大,更接近其他类别。该实验说明不确定性可表征对应特征对数据的重要性,若不确定性很小,则该特征重要程度高,只能接受较小的变化,一旦遭遇较大的改变就可能改变数据的性质;相反若不确定性较大,则说明该特征重要性较低,可以承受较大的变化而不改变数据的性质。这个现象与向量数据实验中精度提高的原因分析一致。以上是特征级别上不确定性的意义,下面分析神经元级别不确定性的意义。一种直观的想法是:如果某神经元对应相同类别的数据,那么该神经元的不确定性较小;相反若神经元对应不同类别的数据,那么该神经元的不确定性较大。为了验证这一点,进行图4的实验,即重建训练好的网络权向量,并绘制神经元节点方差 $L_2$ 范数的热图。可见GNSOM同样具有保持数据拓扑结构不变的特点,相同类别的数据会聚集在相邻的区域。其次从方差 $L_2$ 范数的热图中可见不同类别分界线上神经元的方差比周围神经元的方差略大。这一结果符合直觉,因为在分界线上意味着该神经元会受到不同类别数据的影响,不确定性理应偏大。这一特点在训练结束后对于分析无监督学习的结果有重要的参考价值,因为方差比周围大的神经元可看作是分界线上的神经元,这有助于确定数据类别。此外,这是所有维度方差 $L_2$ 范数对应的热图,实际上每个特征方差的热图也有类似特点,总体方差的热图结合单个特征方差的热图,可获得更丰富有效的信息,这与SOM相比有明显的优势。

(2)不确定性如何影响模型结果。根据式(8)可知,在模型训练过程中,神经元的方差能学到数据中隐含的不确定性信息,并通过式(7)将方差作为均值的加权项来反映各个特征的重要性。也即不确定性信息会让模型注重于拟合重要程度更高的特征,而对于重要性较低的特征则允许存在一定的偏差,所以GNSOM的神经元能够更好地刻画出对应类别的信息,这就是不确定性对模型产生的影响。

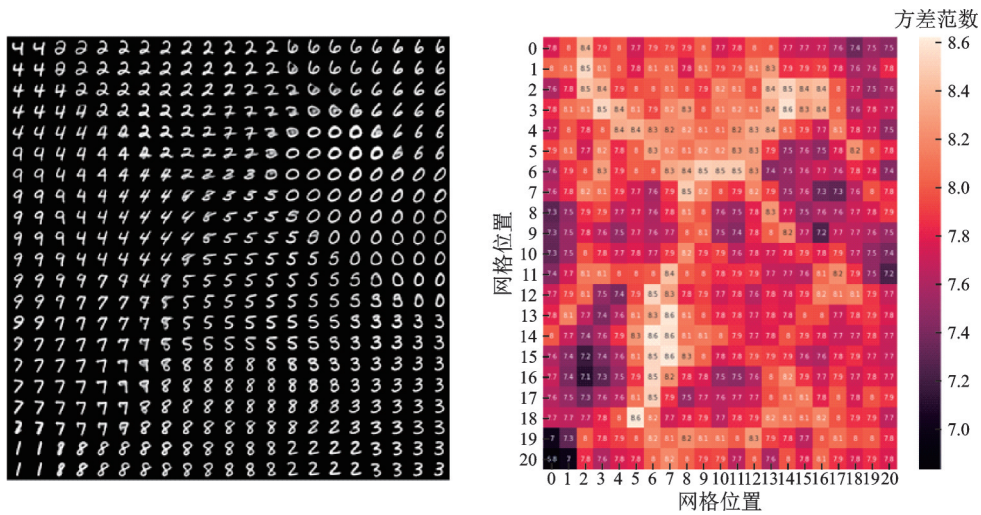


图4 神经元节点图像重建及神经元方差热图

Fig.4 Image reconstruction and variance heat map of neuron nodes

此外,分布型神经元包含的信息更多,较少的分布型神经元就能得到相对更好的训练结果,这可以通过实验验证。如表3所示,随着神经元网格边长逐渐减小,SOM的准确率出现较大的降幅,而GNSOM的准确率则仍能更好地保持。

### 3.4 时间复杂度分析

在学习与更新过程中,GNSOM相较于SOM的主要区别在于GNSOM新增了刻画不确定性的方差项。如果使用完整的协方差矩阵刻画不确定性则复杂度会相当高,故采用对角矩阵来刻画方差。在这种设置下,具体实现时相当于每个神经元仅新增了1个方差向量,故空间复杂度是原来的两倍,而时间复杂度也是原来的常数倍,两种方法在几个数据集上的运行时间如表4所示。

表4 运行时间

Table 4 Operation time

数据集	iris	wine	seeds	pendigits	MNIST	FASHION_MNIST
SOM	0.90	0.52	1.21	25.73	139.55	121.34
GNSOM	2.15	1.69	3.31	122.75	622.88	598.98

## 4 结束语

本文将SOM模型拓展为由高斯化神经元节点形成的GNSOM模型,为其配备了数据不确定性的表述功能,并用JS分布散度代替原SOM的欧氏距离作为新的相似性匹配度量,由此获得了不确定性表示。在多个数据上不仅验证了GNSOM的有效性,而且分析了其中的不确定性意义以及对模型学习的影响。需要指出的是,本文对神经元权向量的高斯化不确定性处理能作为独立的手段应用于其他网络模型。

### 参考文献:

[1] KOHONEN T. The self-organizing map[J]. Proceedings of the IEEE, 1990, 78(9): 1464-1480.  
 [2] JOHNSSON M. Applications of self-organizing maps[M]. London: IntechOpen, 2012.

表3 不同网格边长的比较

Table 3 Comparison of different grid side lengths

网络边长	Size = 21	Size = 18	Size = 15
SOM	88.24±0.47	87.22±0.76	86.08±1.17
GNSOM	90.63±0.64	89.95±0.37	89.42±0.47

- [3] MWASIAGI J. Self organizing maps: Applications and novel algorithm design[M]. London: IntechOpen, 2011.
- [4] MILJKOVIĆ D. Brief review of self-organizing maps[C]//Proceedings of 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). [S.l.]: IEEE, 2017: 1061-1066.
- [5] LÓPEZ-RUBIO E, PALOMO E J. Growing hierarchical probabilistic self-organizing graphs[J]. *IEEE Transactions on Neural Networks*, 2011, 22(7): 997-1008.
- [6] GHARAEI Z, GÄRDENFORS P, JOHNSON M. Action recognition online with hierarchical self-organizing maps[C]//Proceedings of 2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS). [S.l.]: IEEE, 2016: 538-544.
- [7] FURAO S, HASEGAWA O. An incremental network for on-line unsupervised classification and topology learning[J]. *Neural Networks*, 2006, 19(1): 90-106.
- [8] XING Y, SHI X, SHEN F, et al. A self-organizing incremental neural network based on local distribution learning[J]. *Neural Networks*, 2016, 84: 143-160.
- [9] VAIDYA H, DESELL T, ORORBIA A G. Reducing catastrophic forgetting in self organizing maps with internally-induced generative replay[C]//Proceedings of the AAAI Conference on Artificial Intelligence. [S.l.]: AAAI, 2022: 13069-13070.
- [10] LI N, JIANG K, MA Z, et al. Anomaly detection via self-organizing map[C]//Proceedings of 2021 IEEE International Conference on Image Processing (ICIP). [S.l.]: IEEE, 2021: 974-978.
- [11] FOREST F, LEBBAH M, AZZAG H, et al. Deep embedded self-organizing maps for joint representation learning and topology-preserving clustering[J]. *Neural Computing and Applications*, 2021, 33(24): 17439-17469.
- [12] D'URSO P, DE GIOVANNI L, MASSARI R. Smoothed self-organizing map for robust clustering[J]. *Information Sciences*, 2020, 512: 381-401.
- [13] BASSANI H F, ARAUJO A F R. Dimension selective self-organizing maps with time-varying structure for subspace and projected clustering[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, 26(3): 458-471.
- [14] WANG X, IMURA J. A Gaussian process-based incremental neural network for online clustering[C]//Proceedings of 2019 IEEE International Conference on Smart Cloud (SmartCloud). [S.l.]: IEEE Computer Society, 2019: 143-148.
- [15] KHACEF L, MIRAMOND B, BARRIENTOS D, et al. Self-organizing neurons: toward brain-inspired unsupervised learning [C]//Proceedings of 2019 International Joint Conference on Neural Networks (IJCNN). [S.l.]: IEEE, 2019: 1-9.
- [16] ARAÚJO A F R, ANTONINO V O, PONCE-GUEVARA K L. Self-organizing subspace clustering for high-dimensional and multi-view data[J]. *Neural Networks*, 2020, 130: 253-268.
- [17] ANDRAS P. Kernel-Kohonen networks[J]. *International Journal of Neural Systems*, 2002, 12(2): 117-135.
- [18] KENDALL A, GAL Y. What uncertainties do we need in Bayesian deep learning for computer vision?[J]. *Advances in Neural Information Processing Systems*, 2017, 30: 5574-5584.
- [19] SHI Y, JAIN A K. Probabilistic face embeddings[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2019: 6902-6911.
- [20] CHANG J, LAN Z, CHENG C, et al. Data uncertainty learning in face recognition[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2020: 5710-5719.
- [21] YOU S, XU C, XU C. Online dictionary learning with confidence[C]//Proceedings of 2018 IEEE International Conference on Data Mining (ICDM). [S.l.]: IEEE, 2018: 707-716.
- [22] ENDRES D M, SCHINDELIN J E. A new metric for probability distributions[J]. *IEEE Transactions on Information Theory*, 2003, 49(7): 1858-1860.

#### 作者简介:



刘达(1997-),男,硕士研究生,研究方向:模式识别、机器学习, E-mail: liuda-da@nuaa.edu.cn。



陈松灿(1962-),通信作者,男,教授,研究方向:模式识别、图像处理、神经网络理论及应用, E-mail: s.chen@nuaa.edu.cn。