

# 具有生物真实性的强抗噪性神经元激活函数

麦应潮 陈云华 张 灵

(广东工业大学计算机学院 广州 510006)

**摘要** 当前人工神经网络虽然在图像识别等方面媲美人脑,但因其所采用的激活函数 ReLU 和 Softplus 等只是对生物神经元输出响应特性的高度简化与模拟,使其在抗噪性、不确定性信息处理及功耗等方面与人脑仍存在巨大差距。通过分析生物神经元仿真实验,以其响应特性为基础,引入反映每个神经元随机性的参数  $\eta$ ,构建出一种具有生物真实性的强抗噪性激活函数 Rand Softplus。最后将该激活函数应用于深度残差网络,并基于人脸表情数据集对其进行验证。结果表明,在输入无噪声或具有少量噪声时,文中提出的激活函数与当前主流激活函数的识别精度基本持平,当输入包含较大噪声时,文中所提激活函数的识别精度远高于其他激活函数,表现出了良好的抗噪性能。

**关键词** 激活函数,抗噪性,LIF 模型,神经网络

中图分类号 TP183 文献标识码 A DOI 10.11896/j.issn.1002-137X.2019.07.031

## Bio-inspired Activation Function with Strong Anti-noise Ability

MAI Ying-chao CHEN Yun-hua ZHANG Ling

(School of Computers,Guangdong University of Technology,Guangzhou 5110006,China)

**Abstract** Although the artificial neural network is almost comparable to the human brain in image recognition,the activation functions such as ReLU and Softplus are only highly simplified and simulated for the output response characteristics of biological neurons. There is still a huge gap between the artificial neural network and the human brain in many aspects,such as noise resistance,uncertainty information processing and power consumption. In this paper,based on the simulation experiments of biological neurons and their response characteristics,a strong anti-noise activation function Rand Softplus with biological authenticity was constructed by defining and calculating parameters  $\eta$ ,which reflects the randomness of each neuron. Finally,the activation function was applied to the depth residuals network and verified by facial expression dataset. The results show that the recognition accuracy of the activation function proposed in this paper is almost equal to the current mainstream activation function when there is no noise or a small amount of noise,and when the input contains a large amount of noise,it shows good anti-noise performance.

**Keywords** Activation function, Anti-noise,Leaky integrate-and-fire model,Neural networks

## 1 引言

人工神经网络<sup>[1-2]</sup>通过模仿人脑连接结构与刺激响应机制来处理外部数据,在多个领域的研究中取得了突破性进展<sup>[3-4]</sup>。尽管如此,人工神经元模型与生物神经元模型之间仍存在巨大差异,前者是对后者响应特性的高度简化,缺乏生物真实性。

迄今为止,已构建出的几种典型生物神经元模型包括 Hodgkin-Huxley<sup>[5]</sup>模型、LIF<sup>[6]</sup>(Leaky Integrate-And-Fire)模型、IF(Integrate-And-Fire)模型、Izhikevich<sup>[7]</sup>模型等。其中,Hodgkin-Huxley 模型最接近真实生物神经元,但其由于计算复杂度高、参数优化困难<sup>[8]</sup>,难以得到实际应用。LIF 模型不

仅能很好地描述生物神经元的响应机制,而且具有较低的计算复杂度,因此成了最广泛使用的生物神经元模型。上述生物神经元模型能够准确地描述神经元细胞的输入电流与脉冲发放之间的非线性关系,具有高度的生物真实性,但同时也具有计算复杂度过高的缺点。

人工神经元较之生物神经元,丢失了信息中的时间属性,仅保留了空间属性,通过激活函数对生物神经元输入与输出的响应情况进行高度的抽象与简化。由于人工神经元的输入与输出已不具备生物神经元的物理量纲,其激活函数只是对生物神经元模型输入与输出间非线性关系的一种模拟。早期常使用 Sigmoid 和 Tanh 作为激活函数。上述函数均存在梯度消失问题<sup>[9]</sup>,限制了神经网络的深度。为解决该问题,研究

到稿日期:2018-05-22 返修日期:2018-09-16 本文受广东省自然科学基金(2016A030313713),广东省交通运输厅科技项目(科技-2016-02-030),广东省科技计划项目(2013B040500008)资助。

麦应潮(1994-),男,硕士,主要研究方向为神经形态计算、机器学习;陈云华(1977-),女,博士,副教授,CCF 会员,主要研究方向为深度学习、神经形态计算,E-mail:yhchen@gdut.edu.cn(通信作者);张 灵(1968-),女,博士,教授,主要研究方向为模式识别、智能化信息处理、人工智能等。

者提出了近似生物神经元响应的激活函数——ReLU<sup>[10]</sup>和Softplus。与ReLU相比,Softplus函数更接近于生物神经元的响应曲线,同时梯度的平滑和非零特性使得其训练得到的网络具有更好的稳定性。但非零特性也会导致网络缺乏稀疏性,这与已知的人类大脑对信息的高度稀疏表示相违背。而ReLU函数能产生具有绝对零的稀疏表示,更符合人脑的稀疏特性。其在反向传播的过程中可以很好地避免梯度消失问题,使得深度网络不再需要进行无监督预训练<sup>[11]</sup>。研究表明,使用ReLU的深层网络比使用Sigmoid或Tanh的网络更容易收敛。ReLU由于其简单性和有效性,已成为深度学习领域最广泛使用的激活函数。但ReLU也存在缺点,其非负特性将会导致激活后的数值平均数大于0,产生均值漂移现象<sup>[12]</sup>。ReLU和Softplus仍然只是对生物神经元响应特性的一种近似。因此,人工神经网络虽然在图像识别等方面媲美人脑,但在抗噪性、不确定性信息处理及功耗等方面与人脑仍存在巨大差距<sup>[13]</sup>。

为提高人工神经网络的抗噪性,缩小其与人脑的差距,本文通过分析LIF神经元仿真实验结果,确定其输入输出响应特性曲线,并将其与人工神经网络激活函数曲线进行对比分析,通过定义和计算反映各个神经元随机性的参数 $\eta$ ,并将其作为比例系数以构建一种具有生物真实性的强抗噪性激活函数Rand Softplus。

实验结果表明,当输入具有少量噪声时,网络的识别精度仍能媲美当前主流激活函数,而在输入具有较大噪声时,本文提出的激活函数展现出了更强的抗噪能力。

## 2 LIF生物神经元模型

### 2.1 LIF神经元的理论输出响应特性

LIF神经元模型遵循以下膜电位动力学:

$$\tau_m \frac{dV}{dt} = V_{rest} - V + R_m I(t) \quad (1)$$

其中, $V_{rest}$ 表示神经元的静息电位, $t$ 表示时间,瞬时膜电位 $V$ 随输入电流 $I$ 的变化而变化,膜时间常数 $\tau_m = R_m * C_m$ ,其中 $R_m$ 为膜电阻, $C_m$ 为膜电容。如果给定一个恒定电流 $I$ ,则LIF神经元的响应函数,即发放频率为:

$$\lambda_{out} = [t_{ref} - \tau_m \log(1 - \frac{V_{th} - V_{rest}}{IR_m})]^{-1} \quad (2)$$

其中, $t_{ref}$ 表示绝对不应期,在绝对不应期中所有输入都将被忽略, $V_{th}$ 表示脉冲发放阈值,当 $IR_m > V_{th} - V_{rest}$ 时,会触发一个脉冲,否则膜电位达不到阈值 $V_{th}$ ,神经元发放频率为零。在实际情况下,输入的电流并不是恒定电流,而是带有噪声的电流,LIF神经元对噪声电流的响应函数如下:

$$\lambda_{out} = \left[ t_{ref} + \tau_m \int_{\frac{V_{rest} - V_{th}}{\sigma \sqrt{\tau_m}}}^{\frac{V_{th} - V_{rest}}{\sigma \sqrt{\tau_m}}} \sqrt{\pi} \exp(u^2) (1 + \text{erf}(u)) du \right]^{-1} \quad (3)$$

其中, $\mu, \sigma^2$ 分别表示输入噪声电流的均值与方差, $\text{erf}(u)$ 表示误差函数。

图1表示的是不同标准差的噪声电流输入下LIF神经元的响应曲线<sup>[14]</sup>,由式(3)计算得到。ReLU激活函数的提出正是受到噪声标准差为零时神经元激活曲线的启发。从图1可以看出,随着输入电流噪声标准差的增加,LIF神经元响应曲

线将逐渐接近于Softplus激活函数的形状。由此可见,受噪声电流的影响,生物神经元的输出响应区域大致位于ReLU与Softplus之间。

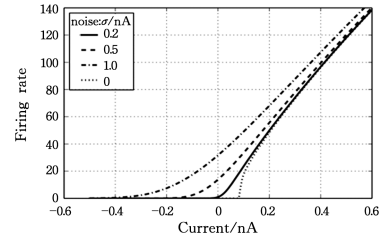


图1 输入不同标准差的噪声电流下LIF神经元的响应曲线  
Fig. 1 Response curves of LIF neurons under currents with different standard deviations of noise

### 2.2 噪声电流的表示形式

本文通过泊松脉冲序列来模拟真实情况的噪声电流<sup>[15]</sup>,并将其作为仿真实验的输入。噪声电流的均值 $\mu$ 和方差 $\sigma^2$ 如下:

$$\mu = \tau_{syn} \sum_i w_i \lambda_i \quad (4)$$

$$\sigma^2 = \frac{1}{2} \tau_{syn} \sum_i w_i^2 \lambda_i \quad (5)$$

其中, $i$ 表示第 $i$ 个神经元, $\tau_{syn}$ 表示突触时间常数,神经元之间以 $w_i$ 的权重相连接,用 $\lambda_i$ 表示神经元发放率。

选用NEST<sup>[16]</sup>和PyNN<sup>[17]</sup>作为仿真平台,对LIF神经元的响应情况进行仿真<sup>[14]</sup>。实验中脉冲序列分别通过一个兴奋性和抑制性突触输入到LIF神经元中以模拟噪声电流。两个输入脉冲频率由均值 $\mu$ 和方差 $\sigma^2$ 决定,通过式(4)和式(5)计算得到。仿真过程中,将具有特定均值 $\mu$ 和方差 $\sigma^2$ 的脉冲序列输入LIF神经元中,记录LIF神经元的脉冲发放率,考虑到生物神经元的峰值过程具有随机性,将重复10次实验,并对脉冲发放率取平均值。最后将该平均值与噪声电流驱动下的神经元脉冲发放率做对比。

图2给出了不同驱动方式下LIF神经元的响应情况<sup>[14]</sup>。横坐标表示输入电流的均值 $\mu$ ,纵坐标表示当前输入下的神经元脉冲发放率。噪声标准差 $\sigma$ 选用0.2,0.5,1.0为代表。

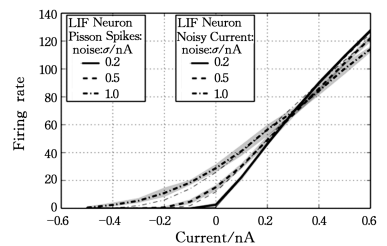


图2 不同驱动方式下LIF神经元脉冲发放率  
Fig. 2 Response firing rate of LIF neuron based on different driving methods

实验表明,当噪声标准差确定时,采用泊松脉冲驱动的LIF神经元响应曲线接近于噪声电流驱动下的响应曲线。这表明通过特定频率泊松脉冲序列模拟噪声电流是可行的,同时生物神经元之间的信息传递是基于脉冲的,通过脉冲序列来模拟输入电流更具有生物真实性。因此,在本文的后续实验

中,采用具有特定均值 $\mu$ 和方差 $\sigma^2$ 的泊松脉冲序列来表示输入噪声电流。

### 2.3 LIF神经元的仿真与分析

图2不仅证明了利用泊松脉冲序列模拟真实噪声电流方法的可行性,同时也表明LIF神经元的响应曲线与噪声方差 $\sigma^2$ 有关。真实情况下,噪声电流具有不确定性,为了更好地确定LIF神经元在真实情况下的响应特性,实验将对噪声标准差 $\sigma$ 以0.1的间隔在0~1取值。在仿真实验中,将泊松脉冲序列驱动的具有均值 $\mu$ 和方差 $\sigma^2$ 的噪声电流注入LIF神经元中,记录LIF神经元的脉冲发放率。为确保拥有足够数据量,实验重复进行5次。

实验结果如图3所示,图中灰点表示具有特定均值 $\mu$ 和方差 $\sigma^2$ 的噪声电流下的脉冲发放率。黑色加粗边界表示噪声电流输入下的神经元响应区域。通过分析响应区域的形状可知,其上边界对应于Softplus函数形状,下边界对应于ReLU函数形状,噪声电流作用下的神经元响应区域介于两者之间。可见,生物神经元响应特性与当前激活函数存在着巨大差异,即ReLU和Softplus激活函数只能用于近似标准差为0或1的噪声电流作用下的生物神经元响应特性曲线,丢失了生物神经元响应过程中的随机性。

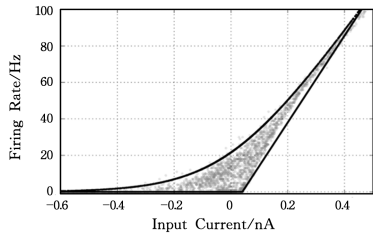


图3 LIF神经元激活范围

Fig. 3 LIF neuronal activation area

## 3 Rand Softplus 激活函数

为了提高人工神经网络的抗噪性能,缩小其与生物神经网络之间的性能差距,本文提出一种能对输入噪声数据进行自适应处理的激活函数:

$$y = (1 - \eta) \max(0, x) + \eta \log(1 + e^x) \quad (6)$$

其中, $x$ 表示神经元输入值, $\eta$ 表示比例系数,该系数与输入的噪声水平有关。

图4给出了Rand Softplus函数的激活情况,每条曲线表示在特定 $x$ 和 $\eta$ 的作用下函数的输出。

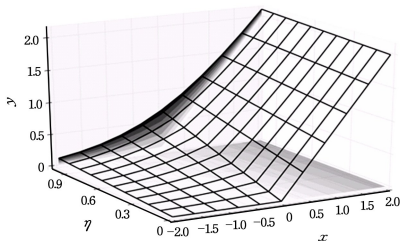


图4 Rand Softplus函数曲线

Fig. 4 Rand Softplus function curve

图4中, $x$ - $y$ 平面投影表示函数输出区域,与图3所示的响应区域相对应,投影颜色表示 $\eta$ 值大小,而 $x$ - $\eta$ 平面的投影

颜色则对应函数输出大小,其颜色越深,数值就越大。其导数定义如下:

$$y'(x) = \begin{cases} \frac{\eta e^x}{1 + e^x}, & x \leq 0 \\ 1 - \eta + \frac{\eta e^x}{1 + e^x}, & x > 0 \end{cases} \quad (7)$$

Rand Softplus函数的导数受比例系数 $\eta$ 的影响。当 $\eta$ 越趋向0,正值输入部分的导数越趋向于1,当输入负值时,梯度随 $\eta$ 的增大而增大,在反向传播中避免梯度的消失。

Rand Softplus激活函数的关键在于比例系数 $\eta$ 的设置,将其与ReLU与Softplus有机地结合起来,构建了一个能描述生物神经元随机响应特性的激活函数。比例系数 $\eta$ 的计算步骤如下:

(1)计算输入数据所蕴含噪声的标准差

计算神经元当前输入数据中的蕴含噪声的标准差。在脉冲神经网络中,神经元 $j$ 的输入电流值可表示为:

$$\mu_j = \sum_i w_{ij} (\lambda_i \tau_{syn}) \quad (8)$$

当其转换成人工神经网络后,神经元的输入表示为:

$$x_j = \sum_i w_{ij} x_i, \text{ and } x_i = \lambda_i \tau_{syn} \quad (9)$$

人工神经网络中输入的噪声方差可由式(10)表示:

$$\sigma_j^2 = \sum_i \frac{1}{2} (w_{ij}^2) x_i \quad (10)$$

对式(10)进行开方运算,即可计算得到人工神经网络输入数据蕴含的噪声标准差。

(2)调整噪声标准差

实验表明,在同层网络中,部分标准差会远大于该层的均值,而随着网络层次的加深,噪声标准差呈逐渐减小的趋势。表1列出了深层网络中噪声标准差变化的趋势,其中噪声标准差取多次实验的均值。为了解决部分标准差远离均值的问题,避免下一步归一化后造成数据分布不均匀,本文重新对大于2倍均值的噪声标准差进行设定,将大于2倍均值的标准差值设定为2倍均值。

表1 噪声标准差均值

Table 1 Mean standard deviation of noise

网络层	第一层	第三层	第五层	第七层
噪声标准差	0.4~1.25	0.21±0.01	0.085±0.01	0.04±0.01

(3)噪声标准差归一化处理

为了解决噪声标准差随网络层次加深而减小的问题,本文将调整后的标准差值进行归一化处理,将其归一化到0~1的范围内,得到比例系数 $\eta$ 。这样有效地解决了随着标准差趋近于0,本文提出的激活函数趋近于ReLU,使得激活函数丧失随机性的问题。

(4)比例系数 $\eta$ 的置零操作

比例系数阈值 $\eta_m$ 用于对网络的稀疏性进行控制,本文将归一化后得到的 $\eta$ 值中小于比例系数阈值 $\eta_m$ 的部分设置为0。

## 4 实验以及结果分析

### 4.1 实验环境及数据集

本文实验采用了Pytorch深度学习框架,使用Python编

程语言,在 Ubuntu14.04 操作系统上进行。硬件平台:i7 6700k CPU,Nvidia Geforce GTX1080 GPU,内存 32 GB。采用下述数据集来评估本文提出的激活函数的性能。图 5 为图像数据样例。

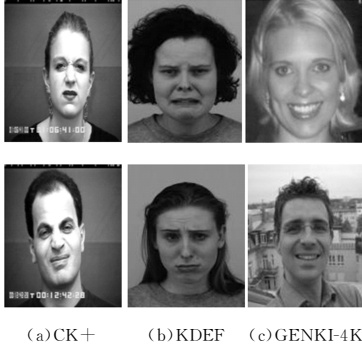


图 5 实验数据集样例(1)

Fig. 5 Sample (1) of experimental dataset

CK+数据集(The Extended Cohn-Kande Dataset)<sup>[18]</sup> 包含了 123 个人的 593 个表情序列,其中带标签的有 309 个表情序列。本实验从数据集中筛选出带有标签的表情序列,从每个序列中随机选取图片并按照标签类别分成 7 类,以 4:1 的比例构建训练集和测试集。

KDEF(Karolinska Directed Emotional Faces) 数据集<sup>[19]</sup> 包含了 70 个人的 7 种不同表情,共有 4900 张图片。由于数据集中每个角度拍摄两张图片,本文从每个角度选取 1 张图片,共 2800 张图片用于构建数据集,该数据集中的图片按照表情分为 7 类,随机选取每个人不同表情的一种姿态为测试集,其余图片作为训练集,以 4:1 的比例构建训练集与测试集。

GENKI-4K 数据集<sup>[20]</sup> 包含 4000 张人脸照片,这些照片分为微笑和不微笑两类,涵盖了不同地区、光照、角度、成像条件以及不同的头部姿态的人脸数据图像,两个类别分别按照 3:1 的比例构建训练集和测试集。

#### 4.2 网络的训练

为了加快网络的收敛速度,本文提出了以下 3 种训练方案:

(1)使用 ReLU 函数作为激活函数来对网络进行初步训练,当网络基本收敛时,再使用 Rand Softplus 替代 ReLU 函数进行网络训练。

(2)在网络训练的初期,采用  $y = 0.5 * \max(0, x) + 0.5 \log(1 + e^x)$  形式的激活函数来对网络进行训练,当网络基本收敛时,再使用 Rand Softplus 函数替代上述函数进行网络训练。

(3)使用 Softplus 函数作为激活函数来对网络进行初步训练,当网络基本收敛时,再使用 Rand Softplus 替代 Softplus 函数进行网络训练。

方案(2)中,通过直接将比例系数  $\eta$  强制设置为 0.5,有效简化了 Rand Softplus 比例系统的计算,同时在转换成 Rand Softplus 函数后不会造成网络性能的较大下降。方案(1)与方案(3)的区别在于网络初期训练选取的激活函数种类。由于 ReLU 函数与本文提出的 Rand Softplus 函数都具

有稀疏的特性,因此采用方案(1)训练得到的网络具有更高的稀疏度。而 Softplus 函数更接近生物神经元的响应曲线,同时梯度的平滑和非零特性使得网络具有更高的稳定性。

本文对上述 3 种训练方案进行了性能测试,实验采用 ResNet-19<sup>[21]</sup> 残差网络结构和 GENKI-4K 表情数据集。网络训练初期采用不同方案预设的激活函数,当迭代 35 次后,网络达到基本收敛状态,再转用 Rand Softplus 激活函数对网络进行继续训练。实验结果如图 6 所示。实验结果表明,方案(1)和方案(2)在迭代 60 次之后基本能达到等同的识别精度,且明显优于方案(3)使用 Softplus 函数进行预训练。与方案(2)相比,虽然方案(1)在转换激活函数后网络性能明显下降,但它们之间的差距在经过 1~2 次迭代后能迅速缩小,在之后的训练中,方案(1)的识别精度略优于方案(2),同时因为方案(1)使用了 ReLU 函数进行训练,在计算速度上也优于方案(2)。在本文的后续实验中皆采用方案(1)对网络进行训练。

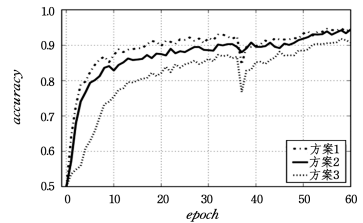


图 6 不同方案识别精度的对比

Fig. 6 Comparison of recognition accuracy of different methods

#### 4.3 网络抗噪性能测试

Rand Softplus 激活函数的提出是为了增强人工神经网络的抗噪性能,本文采用 Resnet-19<sup>[21]</sup> 的网络结构对 3 种表情数据集进行多次实验以验证激活函数种类对网络抗噪性能的影响。实验中对数据集中的数据进行了归一化处理,比例系数阈值  $\eta_m$  取 0.25,数据集的训练集皆不加入噪声,而测试集分别加入均值为 0 但方差不同的高斯噪声。图 7 给出了 Genki-4k 数据集加入零均值,特定方差  $s^2$  的高斯噪声后的图像。3 种表情数据集在不同方差的高斯噪声下的识别精度如表 2 所列。实验结果表明:在没有加入高斯噪声或加入少量高斯噪声的情况下,使用 ReLU 与 Rand Softplus 激活函数的网络识别精度相近,且略优于使用 Softplus 激活函数。当测试集加入方差大于 0.1 的高斯噪声之后,采用 Rand Softplus 激活函数的网络抗噪能力明显优于其他种类的激活函数。通过为激活函数引入随机性来增强网络的抗噪性的方法是可行的。当输入不含噪声或含有少量噪声时,其识别精度仍能媲美当前主流激活函数,这表明随机性的引入并不会对网络的性能造成损失。而当输入具有较大噪声时,具有随机性的 Rand Softplus 激活函数具有更强的抗噪能力。

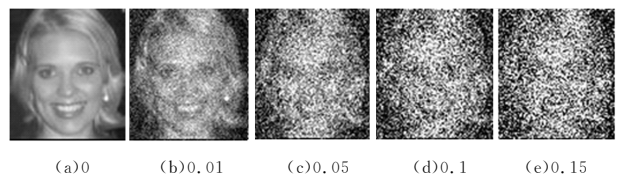


图 7 实验数据集样例(2)

Fig. 7 Sample (2) of experimental dataset

表2 网络抗噪性测试

Table 2 Test of network anti-noise performance

数据集	使用的 激活函数	噪声方差 (0)	噪声方差 (0, 01)	噪声方差 (0, 05)	噪声方差 (0, 1)	噪声方差 (0, 15)
CK+	ReLU	<b>0.9956</b>	0.9934	0.9831	0.9689	0.7840
	Softplus	0.9934	<b>0.9956</b>	0.9333	0.6929	0.5342
	Rand Softplus	<b>0.9956</b>	<b>0.9956</b>	<b>0.9934</b>	<b>0.9825</b>	<b>0.8718</b>
KDEF	ReLU	<b>0.9252</b>	0.9175	0.8755	0.6804	0.4315
	Softplus	0.9036	0.8993	0.8442	0.4820	0.2504
	Rand Softplus	0.9206	<b>0.9184</b>	<b>0.9036</b>	<b>0.7077</b>	<b>0.6204</b>
GENKI-4K	ReLU	0.9432	0.9452	0.9320	0.8805	0.8333
	Softplus	0.9360	0.9463	0.9296	0.8854	0.8100
	Rand Softplus	<b>0.9440</b>	<b>0.9470</b>	<b>0.9360</b>	<b>0.9260</b>	<b>0.8970</b>

**结束语** 本文从生物神经元响应特性出发,引入一个反映生物神经元随机性的参数  $\eta$ ,并将其作为比例系数构建出一种具有抗噪性能的混合激活函数,其中比例系数阈值的设置使得激活函数具有稀疏的特性,也使得训练得到的人工神经网络具有生物大脑数据稀疏表示的特点。相比于其他种类的激活函数,本文提出的 Rand Softplus 激活函数具有更强的生物真实性,在函数响应区域上十分接近生物神经元的响应,一定程度上缩小了人工神经网络与生物神经网络的差距。同时,在零噪声或少量噪声的情况下,本文所提激活函数能达到与 ReLU, Softplus 等常用激活函数基本持平的识别精度,在输入具有较大噪声时,其明显优于其他激活函数。下一步的研究主要是对其计算效率进行优化。

### 参考文献

- [1] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553): 436.
- [2] SCHMIDHUBER J. Deep learning in neural networks: An overview[J]. Neural Networks, 2015, 61: 85-117.
- [3] BADJATIYA P, GUPTA S, GUPTA M, et al. Deep Learning for Hate Speech Detection in Tweets[C]// Proceedings of the 26th International Conference on World Wide Web Companion. Perth: International World Wide Web Conferences Steering Committee, 2017: 759-760.
- [4] KIM B K, LEE H, ROH J, et al. Hierarchical Committee of Deep CNNs with Exponentially-Weighted Decision Fusion for Static Facial Expression Recognition [C]// Proceedings of the 2015 ACM on International Conference on Multimodal Interaction. Seattle: ACM, 2015: 427-434.
- [5] HODGKIN A L, HUXLEY A F. A quantitative description of membrane current and its application to conduction and excitation in nerve[J]. The Journal of Physiology, 1952, 117(4): 500-544.
- [6] STEIN R B. A Theoretical analysis of neuronal variability[J]. Biophysical Journal, 1965, 5(2): 173-194.
- [7] IZHIKEVICH E M. Which model to use for cortical spiking neurons[J]. IEEE Transactions on Neural Networks, 2004, 15(5): 1063-1070.
- [8] ACHARD P, DE SCHUTTER E. Complex parameter landscape for a complex neuron model[J]. PLOS Computational Biology, 2006, 2(7): 794-804.
- [9] GLOROT X, BENGIO Y. Understanding the difficulty of training deep feedforward neural networks[C]// Proceedings of the thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia: PMLR, 2010: 249-256.
- [10] GLOROT X, BORDES A, BENGIO Y. Deep sparse rectifier neural networks[C]// Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics. Fort Lauderdale: PMLR, 2011: 315-323.
- [11] ERHAN D, BENGIO Y, COURVILLE A, et al. Why does unsupervised pre-training help deep learning? [J]. Journal of Machine Learning Research, 2010, 11: 625-660.
- [12] TROTTIER L, GIGU P, CHAIB-DRAA B. Parametric exponential linear unit for deep convolutional neural networks[C]// Machine Learning and Applications. Cancun: IEEE Press, 2017: 207-214.
- [13] STROMATIAS E, NEIL D, PFEIFFER M, et al. Robustness of spiking deep belief networks to noise and reduced bit precision of neuro-inspired hardware platforms[J]. Frontiers in neuroscience, 2015, 9: 222.
- [14] LIU Q, FURBER S. Noisy softplus: a biology inspired activation function[C]// In International Conference on Neural Information Processing. Kyoto: Springer, 2016: 405-412.
- [15] LA C G, GIUGLIANO M, SENN W, et al. The response of cortical neurons to in vivo-like input current: theory and experiment: I. Noisy inputs with stationary statistics [J]. Biological Cybernetics, 2008, 99(5): 279-301.
- [16] GEWALTIG M O, DIEMANN M. Nest (neural simulation tool)[J]. Scholarpedia, 2007, 2(4): 1430.
- [17] DAVISON A P, BRUDERLE D, EPPLER J M, et al. PyNN: a common interface for neuronal network simulators[J]. Frontiers in Neuroinformatics, 2009, 2(11): 204-123.
- [18] LUCEY P, COHN J F, KANADE T, et al. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression[C]// Computer Vision and Pattern Recognition Workshops. San Francisco: IEEE Press 2010: 94-101.
- [19] CALVO M G, LUNDQVIST D. Facial expressions of emotion (KDEF): Identification under different display-duration conditions[J]. Behavior Research Methods, 2008, 40(1): 109-115.
- [20] WHITEHILL J, MOVELLAN J R. A discriminative approach to frame-by-frame head pose tracking[C]// IEEE International Conference on Automatic Face & Gesture Recognition. Amsterdam: IEEE Press, 2008: 1-7.
- [21] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston: IEEE Press, 2016: 770-778.