人工神经网络与生物神经元模型的共生：训练与测试

Tatyana Bogatenko[a]、Konstantin Sergeev[a]、Andrei Slepnev[a]、J¨urgen Kurths[b, c]、Nadezhda Semenova[a]

[a]俄罗斯萨拉托夫国立大学，俄罗斯萨拉托夫市阿斯特拉罕街*83*号，邮编*410012*[b]洪堡大学物理系，德国柏林牛顿街*15*号，邮编*12489*[c]波茨坦气候影响研究所，德国波茨坦特勒格拉芬贝格*A31*号，邮编*14473*

arXiv:2302.01636v1 [cs.NE] 3 Feb 2023

摘要

本文展示了构建由生物神经元数学模型组成的人工神经网络（ANN）的可能性，并探讨了如何识别其特征。我们以菲茨休–南云（FHN）系统为例，说明一种模拟简化神经元活动的模型。首先，为了揭示如何将生物神经元嵌入人工神经网络，我们训练由非线性神经元构成的人工神经网络，以解决一项基础图像识别任务，使用的数据集为MNIST数据库；接着，我们说明如何将FHN系统引入这个训练好的人工神经网络。最后，我们展示了内部包含FHN系统的人工神经网络能够成功训练，且其准确率有所提高。上述工作为模拟神经网络的发展方向带来了广阔前景，在这类网络中，人工神经元可以由生物神经元取代。

关键词： 机器学习，神经网络，FitzHugh–Nagumo 系统，线性回归，人工神经网络，生物神经元

PACS：05.45.-a, 05.10.-a, 47.54.-r, 07.05.Mh, 87.18.Sn, 87.19.ll*MSC*： 70K05, 82C32, 92B20

引言

实现神经网络有两种不同的方法，对神经网络也有两种不同的定义。从非线性动力学的角度来看，神经网络有助于描述生物现象，以及神经回路中神经元之间的相互作用在响应内部或外部影响时所呈现的特征。生物神经元的时间动力学及其彼此之间的连接极其复杂，因此有大量研究描述了不同复杂程度的模型 [1, 2, 3, 4]。

另一方面，还有人工神经网络（ANN）。尽管名称相似，但这些网络在用途和设计上与生物神经网络截然不同。人工神经网络一直是解决众多计算任务的前景广阔且得到广泛应用的工具，应用领域涵盖气候研究、医学、社会学和经济学等各种科学与工程领域 [5, 6]。人工神经网络由人工神经元组成，其作用是根据输入信号的线性或非线性变换生成输出信号。人工神经网络的训练在于拟合并调整其神经元之间的连接矩阵。在学习过程中，连接矩阵以某种方式构建，使网络能够输出所需的结果。这类神经网络的实现理念源于生物学 [7]。20世纪40年代，科学家们受到神经回路构成方式的启发，并尝试实现其简化模型，以解决那些没有严格制定的求解算法的复杂问题，从而解决那些没有严格制定的求解算法的复杂问题 [8]。

人工神经网络最初于20世纪40年代和20世纪50年代开发，此后经历了多次重大改进。第一代简单的阈值神经元只能产生二值输出，后来发展为使用平滑激活函数的系统，从而能够产生实值输出。最新的一类人工神经网络基于脉冲神经元 [9, 10]，称为脉冲神经网络（SNN）。

与前几代神经网络不同，脉冲神经网络会考虑输入信息的时间特性。从这一点来看，脉冲神经网络架构更接近于一种合理的生物神经网络模型，尽管仍经过了高度简化 [11]。在这种网络中，人工神经元之间的信息传递与生物神经元之间的信息传递相似。

与传统的人工神经网络类似，脉冲神经网络按层排列，信号从输入端出发，经过一个或多个隐藏层，传至输出层。不过，脉冲神经网络在隐藏层中使用脉冲神经元，这些神经元由一种现象学模型描述，该模型代表了脉冲生成过程。在生物神经元中，突触前神经元的活动会影响突触后神经元的膜电位；当膜电位越过阈值时，就会产生脉冲 [11, 9]。这一复杂过程已通过多种数学模型加以描述，其中，霍奇金-赫胥黎模型是最早、也最著名的模型 [12]。

电子邮箱地址： semenovani@sgu.ru (Nadezhda Semenova)

为了在计算开销与生物真实性之间取得平衡，人们提出了其他几种模型，例如漏积分发放模型 [13] 或伊日凯维奇模型 [14]。

基于神经科学的现有知识，人们开发了若干种信息编码方法，例如频率编码或时延编码。频率编码通过脉冲的速率（或频率）来解读信息，而时延编码则利用脉冲的时序。这两种方法都是全时序编码的特例。在全时序编码中，时序与某个特定事件相关联，例如参考神经元的一个脉冲。

随着人工神经网络及其任务变得越来越复杂，我们可能很快就会面临某种危机 [15, 16], ：任务变得如此复杂，以至于现代计算机的能力很快将无法满足日益增长的需求。此时，硬件神经网络这一前沿方向便能发挥作用 [17]。按照这种方法，神经网络并非借助计算机创建，而是一种能够学习并解决任务的真实设备。神经元本身及其相互之间的连接存在于物理层面；也就是说，该模型并非在计算机上进行模拟，而是依据其物理原理在硬件中实现。

本研究的主要目的是展示如何创建由生物神经元数学模型构成的训练后神经网络，并识别其特征。本研究以菲茨休-南云系统 [18, 19] 为例。FHN系统是一种广为人知的生物神经元简化模型，在特定条件下能够表现出脉冲动力学。它常用于模拟简化的神经活动。

这项工作有助于拉近人工神经网络与生物神经网络之间的距离，并揭示如何将生物神经元嵌入人工神经网络。也就是说，我们可以采用人工神经网络的拓扑结构（连接方式），以及生物系统的动力学和交互特征。因此，可以使脉冲人工神经网络更接近生物神经网络。

这项工作分为几个阶段。首先，构建一个由人工线性神经元和非线性神经元组成的简单神经网络，并训练它解决一个基础的图像识别问题：识别MNIST数据库中的手写数字（第1节）。然后，在现有的神经网络中引入一定数量的FHN系统网络，任务是确定网络仍能正常工作的条件（第3节）。下一步是增加任务的难度。从这一阶段开始，我们使用一个实现了FHN系统的网络，并尝试训练该神经网络（第4节）。

1. MNIST数据库和网络拓扑结构

首先，训练一个只有一个隐藏层的简单深度神经网络。该神经网络的示意图如图 1 所示。

由于要解决的是手写 MNIST 数字识别任务，人工神经网络的输入信号是由 28×28 个像素组成的图像。通常，这种图像会转换为大小为 1×784 的向量 $\mathbf{X}$，并输入人工神经网络的输入层。因此，第一层由 784 个简单的线性神经元组成，其激活函数为 $f(x) = x$。这些神经元不会变换信号，只会将其传递到下一层。为此，向量 $\mathbf{X}$ 会与相应的连接矩阵$\mathbf{W}^{\text{in}}$ 相乘，该矩阵的大小为 784×100。这样，输入图像就被转换为输入信号，并传递给隐藏层的 100 个神经元。在这一层中，神经元使用 Sigmoid 激活函数：$f(x) = 1/(1 + e^{-x})$。不过，函数的选择不会影响后续结果，使用双曲正切函数也会得到相似的结果。接下来，通过连接矩阵 $\mathbf{W}^{\text{out}}$，将来自 100 个隐藏层神经元的信号传递到输出层。输出层由 10 个神经元组成，它们同样使用线性激活函数 $f(x) = x$。

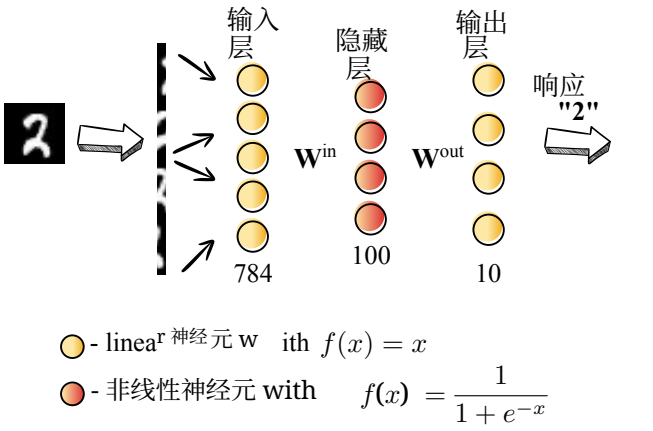


图 1：所研究神经网络的示意图。具有线性激活函数的人工神经元以黄色标示，而具有非线性激活函数、之后将由 FHN系统取代的神经元则以红色标示。

神经网络的响应是输出信号最大的输出神经元的索引编号。此操作也称为 Softmax 函数。具体而言，如果将数字 2 的图像输入人工神经网络，如图 1 所示，那么人工神经网络的响应“2”对应的情况是：编号为 $i = 2$ （其中 $i \in [0;\ 9]$）的神经元具有最大的输出。

人工神经网络使用 MNIST数据库 [20]进行训练。该数据库包含一个训练集（60,000 张数字 0–9 的图像）和一个测试集（10,000 张图像）。训练神经网络时，我们使用了 Keras [21] 库。这是一个免费分发的应用程序接口。训练好的人工神经网络在训练集上的准确率为 99.5%，而在测试集上的准确率为 97.7%。训练的结果是连接矩阵 $\mathbf{W}^{\text{in}}$ 和 $\mathbf{W}^{\text{out}}$。

2. FitzHugh–Nagumo 系统

为了用 FHN系统取代隐藏层中的 100 个人工神经元，需要了解 FHN系统本身的动力学，以及应如何向其馈入输入信号。输入信号$\mathbf{X}$ 乘以

连接矩阵 $\mathbf{W}^{\text{in}}$后，得到一个包含 100 个值的向量。这些值被馈入 100 个神经元的输入端。现在，FHN系统充当隐藏层神经元的角色，每个系统都由以下方程 [18, 19]描述：

$$\begin{aligned} \varepsilon \dot{x} &= x - \frac{x^3}{3} - y \\ \dot{y} &= x + a + I(t), \end{aligned} \tag{1}$$

其中，x 是激活变量，而y 是抑制变量。这是FHN系统一种广泛使用的形式，其中参数 ε 决定时间尺度，$I(t)$是输入信号，而a 是控制参数。根据a的取值，系统会呈现安德罗诺夫–霍普夫分岔：当$|a| > 1$ 时，系统处于“焦点”型稳定平衡态（可激发模式）；当 $|a| < 1$时，该模式称为振荡模式，系统表现出周期性尖峰动力学。

从电流和电压的角度描述系统（1）也很常见。此时，变量x 是类似电压的膜电位，具有三次非线性，可通过正反馈实现再生式自激。变量y 称为恢复变量，具有线性动力学，可提供较慢的负反馈。参数I 对应刺激电流。正电流表示电流从细胞膜外侧流向内侧。

图2展示了系统（1）的相平面。同时，还绘出了相应的激活变量 $\dot{x} = 0$ 和抑制变量 $\dot{y} = 0$ 零流线。激活变量零流线对应于$y = x - x^3/3$直线（图2中的橙色直线），而在没有输入信号时，抑制变量零流线对应于$x = -a$ 直线。对于$a = 1$，两条零流线在$x_0 = -1$、$y_0 = -2/3$处相交。此点是$a > 1$对应的稳定平衡态。

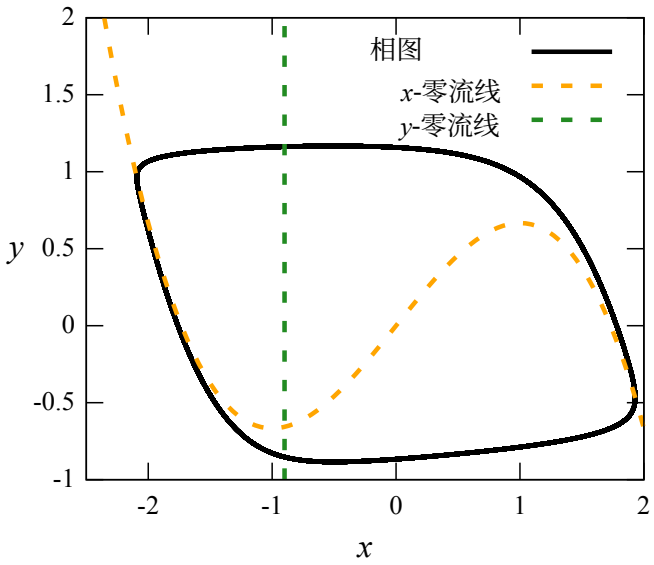


图2：FHN系统（1）的相图，以及相应的 $\dot{x} = 0$（橙色）和 $\dot{y} = 0$ 零流线（绿色）。

如果输入信号不随时间变化，而是向第二个方程中引入额外的恒定分量 $I(t) = I = \text{const}$，则总和 $a + I$ 允许人们影响系统垂直零流线的位置（见图 2，绿色虚线）。因此，由于输入信号 I，系统可以建立振荡模式或可激发模式。对于 $a = 0$ 这些值$|I| < 1$ 将对应于振荡模式，而 $|I| > 1$将建立可激发模式。

3. 将FHN系统嵌入训练好的人工神经网络

由于神经网络最初的训练方式使隐藏层神经元采用“Sigmoid函数”类型的激活函数，因此它们接受取值范围为（$-\infty$；$+\infty$）的输入信号，并输出取值范围为（0；1）的信号。在这种情况下，输入图像向量$\mathbf{X}$ 与连接矩阵$\mathbf{W}^{\text{in}}$ 的乘积可能包含很大的数值，以至于与FHN系统中变量（x，y）的取值尺度不相称。

为了将乘积 $\mathbf{X} \cdot \mathbf{W}^{\text{in}}$ 进一步用作 FHN系统的输入信号，我们提出采用以下归一化方法：

$$I = \gamma \cdot \tanh(\mathbf{X} \cdot \mathbf{W}^{\text{in}}). \tag{2}$$

这样，无论矩阵 $\mathbf{W}^{\text{in}}$ 中的数值有多大，经过双曲正切变换后，数值范围都会被转换为 $(-1;1)$。γ 乘数可用于更精确地设定数值范围。

系统的输出信号定义如下 s:

$$Y = \text{softmax}(\vec{x} \cdot \mathbf{W}^{\text{out}}), \tag{3}$$

与前面一样，这里也使用 Softmax 函数进行计算，但现在它是矩阵 $\mathbf{W}^{\text{out}}$ 与 100 个 FHN系统的变量向量 x_i 的乘积的函数。这使得人工神经网络的响应成为输出信号最大的输出层神经元的索引编号。

图3显示了神经网络对包含数字 0、3 和 5 的三张不同输入图像的响应随时间变化的情况。严格来说，图3并未显示神经网络的即时输出。系统存在一个持续 1000 个无量纲单位 $T^{\text{trans}} = 1000$ 的暂态过程；为避免考虑系统建立过程，这段时间被舍弃。此时，结果的解释存在一个问题。由于 FHN系统是脉冲系统，并且可能呈现振荡模式，人工神经网络的响应也可能振荡，在某些时刻还可能激活“错误”的神经元。为了正确解释人工神经网络的响应，可以选择在人工神经网络输出信号 $T = 100$的整个控制记录中持续时间最长的答案。例如，在图3中，顶部记录中持续时间最长的响应是“0”，中间记录中是“3”，底部记录中是“5”。

其他数字也得到了相同的结果。表1显示了参数 $\gamma = -0.5$ 下训练集和测试集中数字 0–9 的准确率。从表中可以看出，训练集的平均准确率为 73.4%，测试集的平均准确率为 73.3%。

还计算了其他 γ 值下的平均准确率。图4展示了测试集的平均准确率随 γ 变化的情况。从图中可以看出，最低

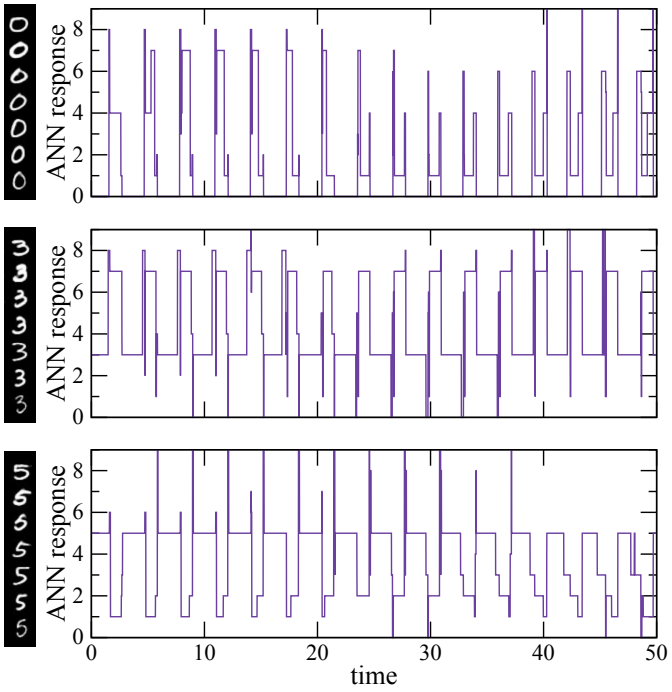


图3：采用已实现的 FHN 系统的人工神经网络输出的时间演化，展示了三种不同输入图像：数字“0”（上图）、“3”（中图）和“5”（下图）。

准确率可在 $\gamma > 0$ 时获得，且 γ 接近于零。此外，对于 $-1 < \gamma < 0$，准确率会随着 γ 的减小而提高，并在 $\gamma < -1$ 时趋于饱和。

4. 人工神经网络训练

为加快训练和测试过程，并未使用训练集和测试集中的全部图像。使用了训练集中的 10000 个样本和测试集中的 1000 个样本，且每种数字的样本数量相同。

FHN 系统的参数保持不变。由于人工神经网络训练涉及大量输入图像和连接矩阵的运行，为加快这一过程，将稳定时间缩短至 $T^{\text{trans}} = 200$，但控制时间 $T = 100$保持不变。重复此前所述步骤并未影响准确率。

人工神经网络训练过程中遇到了一些困难。我们考虑了大量网络拓扑结构和几种非线性激活函数，也不断调整层数。最终，我们得出结论：最优网络拓扑结构如图5所示。将784个像素输入人工神经网络输入，因此输入层仍由784个神经元组成。上一节已说明，在将信号输入FHN之前，必须先使用双曲正切对其进行归一化，因此我们将这一处理步骤加入新人工神经网络的第一层。结果是，第一层的神经元数量保持不变，但其激活函数变为双曲正切。第二层由100个FHN系统组成。输入层通过连接矩阵 $\mathbf{W}^1$ 784×100与第二层相连。

数字	$\gamma = -0.5$		$\gamma = -1$		$\gamma = 0.5$	
	Tr.	Test.	Tr.	Test.	Tr.	Test.
0	99.0	98.6	88.9	99.2	6.7	6.4
1	1.4	1.5	16.3	16.4	0.0	0
2	83.3	82.6	91.8	90.3	5.6	5.6
3	84.6	85.1	92.7	91.9	3.2	3.7
4	90.4	88.0	98.0	94.3	1.0	0.6
5	29.6	31.9	52.0	54.9	0.3	1.1
6	96.9	96.2	99.9	98.0	8.6	9.4
7	56.5	57.5	80.4	78.6	0.9	0.8
8	99.8	99.5	100.0	99.0	6.5	6.0
9	92.3	91.9	98.9	96.6	1.36	1.98
平均值	73.4	73.3	83.0	82.9	3.4	3.6

表1：采用已实现的FHN系统训练（Tr.）和测试（Test.）每种数字类型数据集的人工神经网络准确率。共有三个具有 $\gamma = -0.5$、 $\gamma = -1$和 $\gamma = 0.5$的人工神经网络。

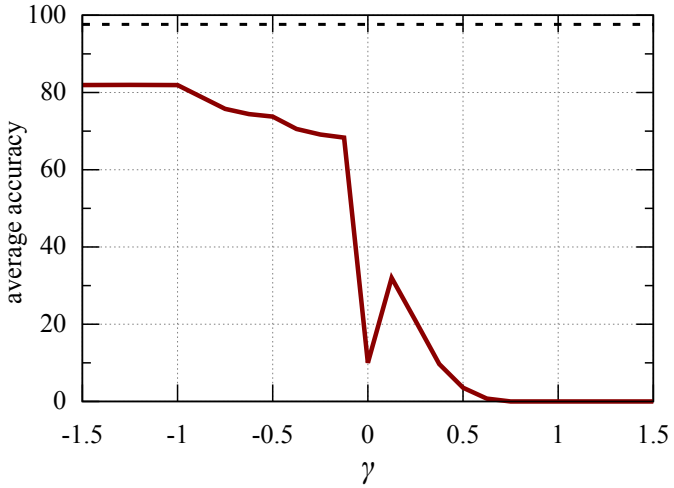


图4：采用已实现的FHN系统的人工神经网络平均测试准确率随参数 γ的变化。

第三层是为了简化学习过程和导数计算而设置的。该层包含100个人工神经元，使用“Sigmoid函数”作为激活函数。第二层和第三层通过恒等连接矩阵$\mathbf{W}^2 = E$相连；该矩阵是固定的，在学习过程中不会发生变化，也就是说，在整个训练过程中，第二层和第三层的神经元始终一一对应。

输出层包含10个线性神经元，其函数为Softmax 函数（）。输出层通过连接矩阵$\mathbf{W}^3$与前一层相连。

对所得人工神经网络输出信号的解释与上一节相同。人工神经网络随时间的响应 $T = 100$ 是产生最大输出信号持续时间最长的神经元的索引编号。

训练采用线性回归的反向传播方法进行。这里应用了以下技巧：在前向传播过程中，使用FHN系统作为第2层神经元；而在反向传播过程中，则将它们替

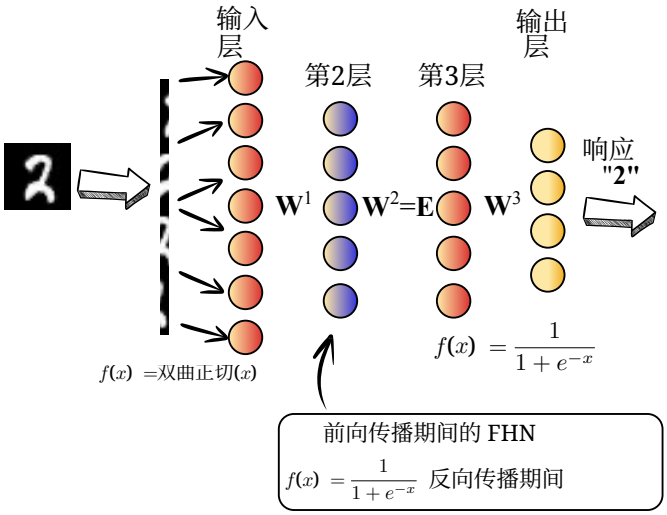


图5：可训练的人工神经网络与 FHN 系统示意图。采用线性激活函数的人工神经元以黄色标示，采用非线性激活函数的神经元以红色标示。紫色表示 FHN 系统。

数字	训练集 (10000)	测试集 (1000)	训练集 (60000)	测试集 (10000)
0	97.1	96.5	95.9	95.6
1	97.5	97.6	96.6	97.8
2	89.2	87.1	86.4	86.6
3	87.1	85.1	84.2	86.8
4	91.9	86.4	88.6	89.1
5	85.9	86.2	81.0	81.7
6	95.2	86.2	92.2	91.1
7	94.4	86.9	92.2	90.5
8	0.1	0	0.1	0.1
9	83.7	80.9	79.5	79.9
平均值	82.2	79.3	79.7	79.9

表 2：采用 FHN 系统的训练好的人工神经网络在各数字类型的训练（Tr.）和测试（Test.）数据集上的准确率。参数 γ 设为 -0.5。

换为带有“Sigmoid函数”类型激活函数的常规人工神经元 $f(x) = 1/(1 + e^{-x})$，以便正确计算导数。图5中这些神经元以蓝色表示。图6中的代价函数(a)以及训练集和测试集上的准确率随训练轮次的变化(b)展示了训练过程。

表 2 给出了截断数据集（图 6 基于这些数据集绘制）和完整MNIST数据集（我们已在这些数据集上应用训练好的网络）中按数字划分的训练集和测试集的相应准确率。上述结果是在 $\gamma = -0.5$ 的取值下获得的。所得网络的平均准确率约为 80%。与嵌入了 FHN系统的训练好的人工神经网络的准确率相比，该 γ 值对应的准确率约为 73%。因此，采用所提出的方法进行训练后，神经网络的准确率得以提高。

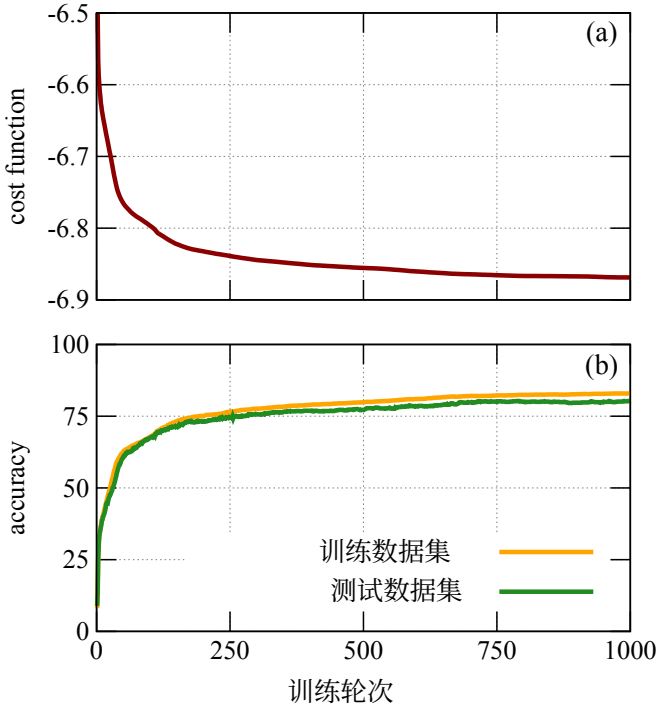


图6：通过训练和测试数据集上的代价函数（a）和准确率（b）展示内部包含 FHN 系统的人工神经网络的训练过程，指标随训练轮次的变化而变化。

不同数字的准确率分布尤其值得关注（见表 2 和附录中的图 A.7）。所得人工神经网络对数字 8 的识别效果不佳。若不考虑数字 8，总体准确率约为 90%。

结论

我们成功找到了将菲茨休-南云系统引入人工神经网络的方法，从而能够将神经网络拓扑结构与菲茨休-南云尖峰动力学的特性结合起来。我们还找到了使所得神经网络能够达到较高准确率的条件。

第一层的信号与相应的连接矩阵相乘，然后根据公式（2）将其发送到人工神经网络隐藏层中的 100 个 FHN 系统。乘数 γ 可以更精确地控制 FHN 输入信号的振幅。我们在此表明，只有负的 γ 值才能达到适当的准确率。

此外，我们提出了一种利用引入的FHN系统训练人工神经网络的方法。其示意图如图6所示。由此得到的神经网络使准确率提高了≈ 12%。

数据可用性

支持本研究结论的数据可向通讯作者提出合理请求获取。

致谢

本研究部分由俄罗斯总统奖学金 SP-749.2022.5 资助。

附录 A. 根据数字类型划分的训练好的人工神经网络测试准确率

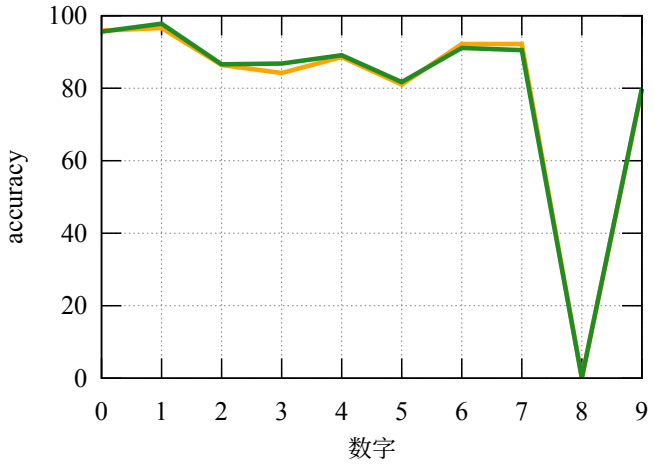


图 A.7：表 2 中的数据按数字类型分组后的测试准确率。

参考文献

[1] B. Hellwig, 《大鼠视觉皮层第 2/3 层锥体神经元局部连接性的定量分析》, 《生物控制论》82 (2) (2000) 111–121。 doi:10.1007/PL00007964。[2] O. Sporns、G. Tononi、G. Edelman, 《连接性与复杂性：神经解剖学与脑动力学之间的关系》, 神经网络 13 (8) (2000) 909–922。 doi:10.1016/S0893-6080(00)00053-8。[3] S. Druckmann、L. Feng、B. Lee、C. Yook、T. Zhao、J. Magee、J. Kim, 《海马区域之间的结构化突触连接》, 神经元 81 (3) (2014) 629–640。 doi:10.1016/j.neuron.2013.11.026。[4] P. Tewarie、B. A. E. Hunt、G. C. O'Neill、A. Byrne、K. Aquino、M. Bauer、K. J. Mullinger、S. Coombes、M. J. Brookes, 《神经元振荡幅度与动态功能连接之间的关系》, 《大脑皮层》29 (6) (2018) 2668–2681。 doi:10.1093/cercor/bhy136。[5] I. Basheer、M. Hajmeer, 《人工神经网络：基础、计算、设计与应用》, 《微生物学方法杂志》 43 1 2000 3 31, doi:10.1016/ ( ) ( ) – 神经计算在微生物学中的应用。 S0167-7012(00)00201-3。[6] O. I. Abiodun、A. Jantan、A. E. Omolara、K. V. Dada、N. A. Mohamed、H. Arshad, 《人工神经网络应用的最新进展：综述》, Heliyon 4 11 2018 00938 doi:10.1016/j.heliyon. ( ) ( ) e 。 2018;e00938。[7] A. Jain、J. Mao、K. Mohiuddin, 《人工神经网络：教程》, 《Computer》29 (3) (1996) 31–44。 doi:10.1109/2.485891。[8] W. S. McCulloch、W. Pitts, 《神经活动中内在观念的逻辑演算》, 《数学生物物理学通报》5 (4) (1943) 115–133。 doi:10.1007/BF02478259。[9] F. Ponulak、A. Kasinski, 《脉冲神经网络简介：信息处理、学习与应用》, 《实验神经生物学学报》71 (4) (2011) 409–433。URL http://europepmc.org/abstract/MED/22237491。[10] A. Tavanaei、M. Ghodrati、S. R. Kheradpisheh、T. Masquelier、A. Maida, 《脉冲神经网络中的深度学习》, 神经网络 111 (2019) 47–63。 doi:10.1016/j.neunet.2018.12.002。

[11] S. Ghosh-Dastidar、H. Adeli, 《脉冲神经网络》, 《国际神经系统杂志》 19 04 2009 295 308, 19731402 doi: ( ) ( ) – 10.1142/S0129065709002002 pMID: 。[12] A. L. Hodgkin、A. F. Huxley, 《膜电流的定量描述及其在神经传导和兴奋中的应用》, 《生理学杂志》 117 4 1952 500 doi:10.1113/jphysiol.1952. ( ) ( ) 。 sp004764 pMID: 。[13] R. Brette、W. Gerstner, 《自适应指数积分发放模型：对神经元活动的有效描述》, 《神经生理学杂志》 94 5 2005 3637 3642, 16014787 doi:10.1152/jn.00686. ( ) ( ) – 2005 pMID: 。[14] E. Izhikevich, 《脉冲神经元的简单模型》, 《IEEE神经网络汇刊》 14 6 2003 1569 1572 doi:10.1109/TNN.2003. ( ) ( ) – 。 820440。[15] J. Hasler、H. Marr, 《探索构建大型神经形态硬件系统的路线图》, 《神经科学前沿》 7 2013 doi:10.3389/ ( ) 。 fnins.2013.00118。[16] S. Gupta、A. Agrawal、K. Gopalakrishnan、P. Narayanan, 《有限数值精度下的深度学习》, 载于：F. Bach、D. Blei（编）, 《第32届国际机器学习会议论文集》, 《机器学习研究论文集》第37卷, PMLR, 法国里尔, 2015年, 第1737–1746页。URL https://proceedings.mlr.press/v37/gupta15.html。[17] G. E. Karniadakis、I. G. Kevrekidis、L. Lu、P. Perdikaris、S. Wang、L. Yang, 《物理信息机器学习》, 《自然综述：物理》3 (6) (2021) 422–440。 doi:10.1038/s42254-021-00314-5。[18] R. FitzHugh, 《神经膜理论模型中的脉冲与生理状态》, 《生物物理学杂志》 1 6 1961 445 466 doi:10. ( ) ( ) – 。 1016/S0006-3495(61)86902-6。[19] J. Nagumo、S. Arimoto、S. Yoshizawa, 《模拟神经轴突的主动脉冲传输线》, 《IRE论文集》50 (10) (1962) 2061–2070。 doi:10.1109/JRPROC.1962.288235。[20] Y. LeCun, 《MNIST手写数字数据库》（1998）。URL http://yann.lecun.com/exdb/mnist/。[21] F. Chollet [ ] 2015 等。Keras 在线版（）。