

基于 DRNet 的常微分方程模型逼近和序列预测方法

李阳^{1,2}, 杨亚莉^{1*}, 钟卫军³

(1. 空军工程大学基础部, 西安, 710051; 2. 文昌航天发射场指挥控制中心, 海南文昌, 571300;
3. 宇航动力学国家重点实验室, 西安, 710043)

摘要 针对残差网络预测精度偏低的问题, 基于系统观测数据和相轨线的关系, 提出密集残差网络的方法实现对自治系统的拟合逼近和序列的高精度预测。首先, 为强化对数据内含“特征信息”的提取和流通使用, 将神经网络各隐藏层的输入与之前各层输出拼接后作为本层的输入, 形成密集连接模块; 其次, 为避免加大网络深度时出现的“退化”现象, 引入残差机制, 将密集连接模块的输入层与输出层相连, 形成密集残差网络。最后, 将密集残差网络应用于线性的单自由度系统振动模型和非线性的 SEIRS 模型、Logistic-Volterra 模型。结果表明, 在规模为 5 000 和 10 000 的数据集上, 密集残差网络对模型的拟合逼近效果和预测精度优于残差网络、反向传播神经网络和密集网络, 特别是在非线性系统上的 4 项定量评价指标均优于对照模型, 表现出密集残差网络对自治系统模型逼近和序列预测的高有效性; 同时, 在观测数据中加入 5% 的噪声后, 密集残差网络表现出较好的抗干扰性。

关键词 深度学习; 密集残差网络; 序列预测; 模型逼近

DOI 10.3969/j.issn.2097-1915.2022.05.013

中图分类号 TP18; O175.1 **文献标志码** A **文章编号** 2097-1915(2022)05-0083-07

DRNet for ODE model Approximation and Series Prediction

LI Yang^{1,2}, YANG Yali^{1*}, ZHONG Weijun³

(1. Fundamentals Department, Air Force Engineering University, Xi'an 710051, China;
2. Command and Control Center, Wenchang Space Launch Site, Wenchang 571300, Hainan, China;
3. State Key Laboratory of Astronautic Dynamics, Xi'an 710043, China)

Abstract In view of the low prediction accuracy of ResNet, a method of DenseResNet (DRNet) for approximation of autonomous systems and series prediction is proposed based on the relationship between observed data and the system phase trajectory. Firstly, in order to strengthen the extraction and the circulation of 'feature information' contained in the data, all the outputs of previous layers in each hidden layer of feedforward neural network are concatenated as an input of this layer to form a dense block. Secondly, to avoid the 'degradation' phenomenon occurs when the depth of neural network increases, the residual mechanism is introduced to connect the input layer and output layer of the dense block to form the DRNet. Finally, DRNet is applied to the linear model, Damped single-degree-of-freedom system and nonlinear models, SEIRS model and Logistic-Volterra model. The results show that DRNet outperforms the Res-

收稿日期: 2022-04-29

基金项目: 国家自然科学基金(11902362); 空军工程大学基础部研究生创新基金

作者简介: 李阳(1989—), 男, 湖北襄阳人, 硕士生, 研究方向为常微分方程及其应用。E-mail: xleon2008@outlook.com

通信作者: 杨亚莉(1974—), 女, 陕西泾阳人, 教授, 硕士生导师, 研究方向为常微分方程及其应用。E-mail: yylhgr@126.com

引用格式: 李阳, 杨亚莉, 钟卫军. 基于 DRNet 的常微分方程模型逼近和序列预测方法[J]. 空军工程大学学报, 2022, 23(5): 83-88. LI Yang, YANG Yali, ZHONG Weijun. DRNet for ODE model Approximation and Series Prediction [J]. Journal of Air Force Engineering University, 2022, 23(5): 83-88.

Net, Back Propagation Neural Network (BPNN) and DenseNet in terms of model approximation and prediction accuracy on both datasets of 5 000 and 10 000. According to the four evaluation indexes on the non-linear models, DRNet has high effectiveness on autonomous systems. The DRNet also shows good noise immunity for its better performance on the data with 5% noise.

Key words deep learning; DenseResNet; series prediction; model approximation; ResNet

基于系统内部机理,常微分方程能够对复杂系统进行准确描述,进而实现对系统状态变量的预测。但是对于内部机理复杂或暂不清楚的系统,直接建立可靠有效的常微分方程较为困难。

近年来,随着软件和硬件的发展进步,对系统状态变量进行观测的方法更加多样高效,可获取的观测数据的精度越来越高、数量越来越多。人们基于观测数据进行了诸多研究。如通过符号回归对系统进行数值建模^[1-2]、无方程建模^[3]、非线性回归^[4]、数据驱动人工神经网络(artificial neural network, ANN)方法^[5]。其中,文献[6~7]使用大量的观测数据驱动深度残差网络对自治系统进行了拟合逼近,实现了对系统状态的预测,但是该方法还存在模型预测精度不够高的问题。

Zhang 等^[8]在深度卷积神经网络的基础上引入残差机制和密集连接机制,提出的残差密集网络实现了图像的超分辨率重建。两种机制的引入使得该网络对数据特征的提取和利用更加充分,同时具有参数规模相对较小、便于收敛、避免“退化”等优点,在图像恢复^[9]、视频超分辨率重建^[10]等领域同样取得了良好效果。受此启发,在不增加观测数据规模的条件下,为进一步提高前馈神经网络(feed-forward neural network, FNN)对自治系统相轨线的拟合逼近效果,本文引入密集连接机制和残差机制,提出使用 DRNet 的方法对自治系统进行拟合逼近,实现对状态变量更高精度的预测。首先,利用系统的观测数据与相轨线的关系,将对自治系统的拟合逼近和序列预测问题转化为使用 FNN 对系统相轨线的拟合逼近问题。其次,为强化深层 FNN 对数据内含“特征信息”的提取和利用,提高对系统的拟合逼近效果,同时避免深层网络的“退化”(degradation)现象,提出并建立 DRNet 网络。而后,选择线性的单自由度系统振动模型^[11]和非线性的 SEIRS 模型^[12]、Logistic-Volterra 模型^[13]为示例,通过模型生成的观测数据训练 DRNet,实现对系统模型的逼近和变量的预测。最后,将 DRNet 对示例模型的预测精度与残差网络(ResNet)、BPNN 和密集网络(DenseNet)进行比较,验证本文 DRNet 方法的有效性,同时将 DRNet 应用于含噪声数据集以验证其抗噪声干扰性。

1 模型建立

本节将对自治系统 S 的模型拟合逼近和序列预测问题转化为使用 FNN 对系统 S 的相轨线进行拟合逼近的问题。

假设自治系统 S 的状态变量 $\mathbf{X}(t)$ 满足方程:

$$\frac{d\mathbf{X}(t)}{dt} = \mathbf{f}(\mathbf{X}(t), \mathbf{p}) \quad (1)$$

式中: $\mathbf{p} = (p_1, p_2, \dots, p_r)^T$, $r \in \mathbf{N}_+$ 为常微分方程的参数, $\mathbf{X}(t) = (x_1(t), x_2(t), \dots, x_q(t))^T$, $q \in \mathbf{N}_+$ 。如图 1 所示,系统 S 状态变量 $\mathbf{X}(t)$ 的相空间为 D ,将方程(1)积分得到系统 S 在 D 内的相轨线^[14]

$$\begin{aligned} \mathbf{X}_l(t) &= \mathbf{X}_l(0) + \int_0^t \mathbf{f}(\mathbf{X}_l(s), \mathbf{p}) ds \\ &\triangleq \mathbf{F}_p(\mathbf{X}_l(0), t) \end{aligned} \quad (2)$$

式中: $\mathbf{X}_l(0)$, $l \in \mathbf{N}_+$ 为系统 S 的初始状态。图 1 中曲线为系统 S 以 $\mathbf{X}(0)$ 为初始状态的相轨线。

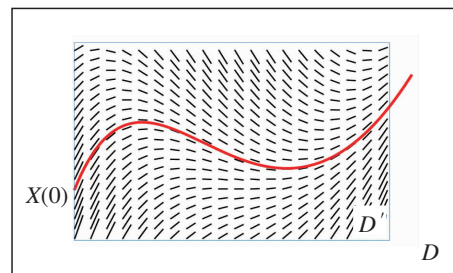


图 1 系统 S 的状态变量在相空间 D 内相图的示意图

假设系统 S 有便于观测的区域,记为 D' 。对于机理复杂或暂不清楚的系统 S,方程式(2)未知。基于 D' 内观测数据对系统 S 的相轨线进行拟合逼近后即可实现对其状态变量的预测。

在 D' 内,对状态变量 $\mathbf{X}(t)$ 进行 n 次独立观测得到:

$$\mathbf{X}^{(1)} = \{\mathbf{X}_j(t_j) \mid j = 1, 2, \dots, n\} \quad (3)$$

由式(2)知, $\{\mathbf{X}_j(t_j) \mid j = 1, 2, \dots, n\}$ 位于 D' 内的 m ($m \leq n$) 条相轨线上。分别以 $\mathbf{X}^{(1)}$ 中的 n 个观测值为初始状态,间隔 Δt 后再次对状态变量 $\mathbf{X}(t)$ 进行观测可以得到集合:

$$\mathbf{X}^{(11)} = \{\mathbf{X}_j(t_j + \Delta t) \mid j = 1, 2, \dots, n\} \quad (4)$$

由集合 $\mathbf{X}^{(1)}$ 和 $\mathbf{X}^{(11)}$ 内状态值的对应关系,可得数据对集合:

$$Data = \{(\mathbf{X}_j(t_j), \mathbf{X}_j(t_j + \Delta t)) \mid j = 1, 2, \dots, n\} \quad (5)$$

在自治系统 S 内, $\mathbf{X}_j(t_j)$ 与 $\mathbf{X}_j(t_j + \Delta t)$ 位于相同的相轨线^[14], 即 $(\mathbf{X}_j(t_j), \mathbf{X}_j(t_j + \Delta t)) = (\mathbf{X}_j(0), \mathbf{X}_j(\Delta t))$,

所以, 式(5)中数据对集合可改写为:

$$Data = \{(\mathbf{X}_j(0), \mathbf{X}_j(\Delta t)) \mid j = 1, 2, \dots, n\} \quad (6)$$

并且由式(2)得:

$$\mathbf{X}_j(\Delta t) = \mathbf{F}_p(\mathbf{X}_j(0), \Delta t), j = 1, 2, \dots, n \quad (7)$$

因此, 对自治系统 S 的研究及状态变量的预测问题可以转化为基于式(6)的观测数据对式(7)中映射 $\mathbf{F}_p(\cdot, \cdot)$ ——即系统 S 相轨线的拟合逼近问题。

根据人工神经网络的万能近似定理^[15], FNN 具有近似任意函数的能力。Wu 等^[16]的研究表明, 当数据对规模相同时, 分布于多条相轨线上的短间隔数据对比分布于少量相轨线上的数据对更能促进 FNN 对系统模型的逼近。因此在构造(6)式的数据集训练 FNN 对系统 S 的相轨线进行拟合逼近时, 优先选择 $\mathbf{X}_j(0)$ 均匀分布于 D' 内的观测数据, 并且观测的时间间隔 Δt 不宜太大。

为训练 FNN 对系统 S 的相轨线进行拟合逼近, 将 $\{\mathbf{X}_j(0) \mid j = 1, 2, \dots, n\}$ 作为 FNN 的输入, 输出记为 $\{\hat{\mathbf{X}}_j \mid j = 1, 2, \dots, n\}$, $\{\mathbf{X}_j(\Delta t) \mid j = 1, 2, \dots, n\}$ 作为 FNN 的目标输出。构造损失函数 L :

$$L = \frac{1}{n} \sum_{j=1}^n \|\hat{\mathbf{X}}_j - \mathbf{X}_j(\Delta t)\|^2$$

式中: $\|\mathbf{X}\|^2$ 表示 \mathbf{X} 的 2 范数的平方。选择合适的 FNN, 使用式(6)的数据集训练 FNN 使得损失函数 L 最小化, 即实现 FNN 对系统 S 的相轨线的拟合逼近。此时, 对系统 S 的相轨线的拟合逼近的问题转化为最优化问题:

$$\min L = \frac{1}{n} \sum_{j=1}^n \|\hat{\mathbf{X}}_j - \mathbf{F}_p(\mathbf{X}_j(0), \Delta t)\|^2 = \frac{1}{n} \sum_{j=1}^n \|\hat{\mathbf{X}}_j - \mathbf{X}_j(\Delta t)\|^2$$

式中: $(\mathbf{X}_j(0), \mathbf{X}_j(\Delta t)) \in Data$ 。

2 DRNet 方法

本节提出并详述 DRNet 方法, 同时建立对照网络模型, 确定模型的定量评价指标。

2.1 DRNet 网络

深度神经网络含有更多的隐藏层, 对函数的逼近效果更好^[17]。Huang 等^[18]在深层卷积神经网络中提出的密集连接机制强化了隐藏层之间特征信息的传输和利用, 同时减少网络参数并提高效率, 在图

像分类识别具有广泛应用。但是过多的隐藏层会使神经网络出现“退化”现象, 即随着网络深度的增加, 模型的准确率增加到一定程度后开始下降^[19]。He 等^[20]提出的 ResNet 极大地消除了深层神经网络的退化现象, 被广泛应用于图像处理的深度学习模型。为进一步提高对自治系统的拟合逼近效果并实现对状态变量更高精度预测, 受残差机制和密集连接机制启发, 本文在具有误差反向传播的 FNN 的基础上提出密集残差网络 DRNet 的方法。

一个 N 层的密集残差网络 DRNet: $\mathbf{R}^n \rightarrow \mathbf{R}^n$, 其网络结构如图 2(a) 所示, 其中“C”表示数据的拼接(具体如图 2(b)所示), “+”表示数据的相加。在 DRNet 中, 第 $i-1$ 层的任意一个神经元均和第 i 层的所有神经元相连(i 表示隐藏层内任意一层的序号)。

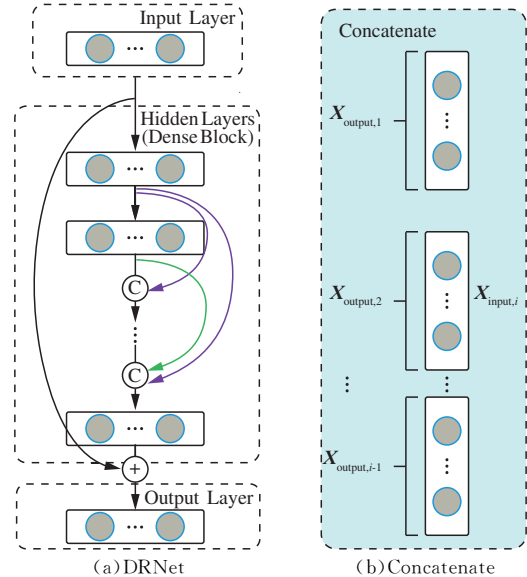


图2 DRNet 网络结构及数据拼接(Concatenate)示意图

在 DRNet 的隐藏层内, 将前 $i-1$ 层的输出拼接后作为第 i 层的输入, 形成密集连接模块。此时, 每层均能直接接收前面各层提取的特征信息, 使得浅层特征和深层特征都得到更加充分的利用, 加强了各隐藏层之间特征信息的流通, 能够提升各层特征信息的利用率, 在一定程度上减少了网络模型的参数数量^[21], 同时, 有效地缓解了梯度消失的问题。因此 DRNet 的第 i 层输入为:

$$\mathbf{X}_{input,i} = [\mathbf{X}_{output,1}, \mathbf{X}_{output,2}, \dots, \mathbf{X}_{output,i-1}]^T \quad (8)$$

式中: $[\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n]$ 表示数据 $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ 的拼接。第 i 层输出为

$$\begin{aligned} \mathbf{X}_{output,i} &= \mathbf{H}_i(\mathbf{X}_{input,i}) = \\ &\mathbf{H}_i([\mathbf{X}_{output,1}, \mathbf{X}_{output,2}, \dots, \mathbf{X}_{output,i-1}]^T) = \\ &\boldsymbol{\sigma}_i(\mathbf{W}_i[\mathbf{X}_{output,1}, \mathbf{X}_{output,2}, \dots, \mathbf{X}_{output,i-1}]^T + \mathbf{b}_i) \end{aligned} \quad (9)$$

式中: $\boldsymbol{\sigma}_i$ 为该层激活函数, 文中均选择为 \tanh ; \mathbf{W}_i 和 \mathbf{b}_i 为该层需要通过训练确定的参数。

为进一步消除深层 DRNet 可能出现的“退化”问题,引入残差机制,将 DRNet 的输入层与输出层直接相连,即第 $N-1$ 层的输出与 DRNet 的输入相加作为 DRNet 的输出:

$$\mathbf{X}_{\text{output}} = \mathbf{X}_{\text{output}, N-1} + \mathbf{X}_{\text{input}} \quad (10)$$

综合考虑系统观测数据的规模、计算消耗、网络收敛性能和误差精度的平衡,本文示例中的 DRNet 网络除输入层和输出层由状态变量 X 的维度决定外,隐藏层数 h 和每层神经元数目 m 分别在 $\{3, 4, 5, 6\}$ 和 $\{20, 30, 40, 50\}$ 中选择。根据 DRNet 在数据集上最小损失函数来确定最优模型的隐藏层数和各层神经元数目。此外,DRNet 的优化器选择为 Adam 算法、损失函数选择为均方误差函数 (Mean Squared Error Loss, MSELoss)、学习率设为 0.01、batch_size 设置为 10。

如图 1 中曲线所示,DRNet 经过观测数据的训练,实现对方程式(2)的拟合逼近后,即可对系统 S 的状态变量进行预测:

$$\begin{cases} \mathbf{X}_k \triangleq \mathbf{X}(k \cdot \Delta t) \\ \text{DRNet}; \mathbf{X}_k \rightarrow \mathbf{X}_{k+1}, k = 0, 1, \dots, T-1 \\ \dot{\mathbf{X}} = \mathbf{X}_T \end{cases} \quad (11)$$

式中:任取初始状态 $\mathbf{X}_0 \in D'$, $\dot{\mathbf{X}} \in D$, $T \in \mathbf{N}_+$, Δt 为单次预测的时间间隔。

2.2 对照网络模型

为评价 DRNet 方法的有效性,本文选择 BPNN、ResNet^[7] 和 DenseNet^[8] 作为对照网络模型。由网络结构可知,DRNet 比 BPNN 多了密集连接机制和残差机制,比 ResNet 多了密集连接机制,比 DenseNet 多了残差机制。对照模型的隐藏层数 h 、每层神经元数目 m 的确定方式、优化器选择和损失函数选择同 DRNet 一致。

2.3 评价指标

为定量评价 DRNet 方法的性能,对于预测值,本文选取的评价指标为均方误差 (mean square error, MSE)、平均绝对误差 (mean absolute error, MAE)、平均绝对百分比误差 (mean absolute percentage error, MAPE)、拟合优度 (R-squared, R^2)。其中, MSE、MAE 和 MAPE 反映出预测值与真实值之间的偏差,其越接近 0,表明预测值越接近真实值。 R^2 反映出预测值与真实值的相似程度,其越接近 1,表示拟合程度越高,预测效果越好,DRNet 对系统模型的拟合逼近越完美。

3 应用结果及分析

为验证 DRNet 方法在常微分方程模型中应用

的有效性,将其分别应用于线性的单自由度系统振动模型和非线性的 SEIRS 传染病模型、Logistic-Volterra 模型。首先,通过式(3)、式(4)获得 $\mathbf{X}^{(1)}$ 和 $\mathbf{X}^{(10)}$ 。不失一般性,假设 $\mathbf{X}^{(1)}$ 在示例模型的便于观测相空间 D' 内服从均匀分布。 $\mathbf{X}^{(10)}$ 通过 LSODE 算子^[22] 基于示例模型方程和 $\mathbf{X}^{(1)}$ 得到。而后,由式(5)、式(6)构造训练 DRNet 和对照网络模型所需的数据集

$$\text{Data} = \{(\mathbf{X}_{\text{input}, j}, \mathbf{X}_{\text{output}, j}) \mid j = 1, 2, \dots, n\} = \{(\mathbf{X}_j, \mathbf{X}_j(\Delta t)) \mid j = 1, 2, \dots, n\}$$

为更加科学地验证 DRNet 的性能,在每个示例模型中,分别构造规模为 5 000 和 10 000 的数据集,记为 Data-5k、Data-10k。此外,为评估 DRNet 的抗噪声性能,在数据集 Data-5k 上加入 5% 的噪声,得到噪声数据集,记为 Data-5kN。

3.1 单自由度系统振动模型

根据受力分析和牛顿第二定律,无驱动有阻尼单自由度弹簧的振动系统的微分方程为:

$$\begin{cases} \frac{dx}{dt} = y \\ \frac{dy}{dt} = -ay - bx \end{cases} \quad (13)$$

对于某一单自由度振动系统,根据模型参数含义,实验中选择 $D' = [-3, 0]^2$,并随机取定参数 $(a, b) = (0.2, 2)$ 。当 $\Delta t = 0.1$ 时,生成数据集,驱动 DRNet 和对照网络模型完成对系统(13)的拟合逼近。根据多次试验的损失函数,在 Data-10k 上表现最优的 DRNet 含有 4 个隐藏层、每层 30 个神经元,在 Data-5k 上表现最优的 DRNet 含有 3 个隐藏层、每层 40 个神经元。对于 D' 内任意初始状态 $\mathbf{X}_0 = (-1, -3)$,训练完毕的最优 DRNet 在 $t \in [0, 20]$ 内的预测结果见图 3。从图中可见,DRNet 在 Data-5k、Data-10k 和 Data-5kN 上均实现了对单自由度系统振动模型的拟合逼近。

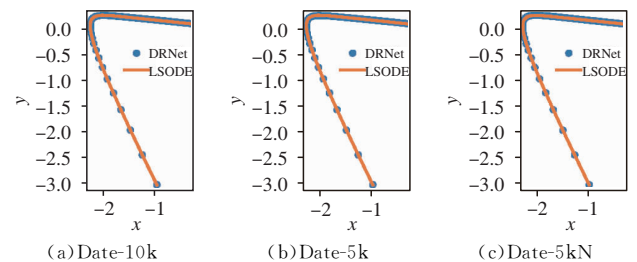


图 3 DRNet 应用于单自由度振动系统模型的结果

DRNet 和对照网络模型的评价指标见图 4,从上往下依次为 Data-10k、Data-5k 和 Data-5kN(以下各图类同)。由图可见,在 Data-5k 和 Data-10k 上,DRNet 的表现均优于对照网络模型。DRNet 在噪声数据集 Data-5kN 上性能稍有下降,除平均绝对

百分比误差 (MAPE) 劣于 ResNet 外 (1.6% > 0.85%), 其他 3 项指标均表现优异。

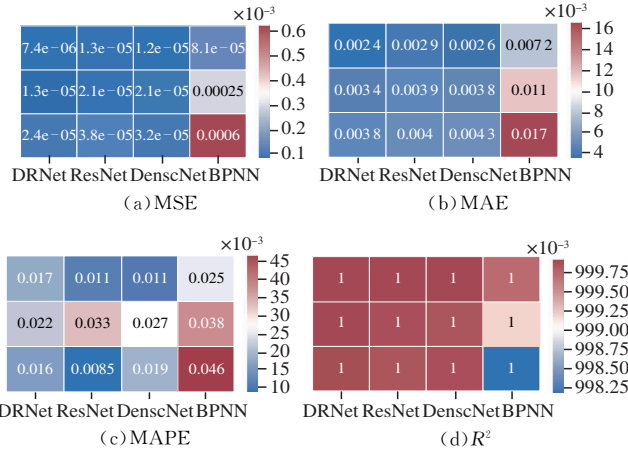


图 4 DRNet 和 ResNet, DenseNet, BPNN 应用于单自由度振动系统模型的评价指标

3.2 SEIRS 模型

利用数学方程(组)模型来描述和分析传染病的传播规律是常微分方程的重要应用方向之一。根据不同传染病的免疫特性,人们陆续建立了不同的仓室模型。如脑膜炎、淋病等,患者康复后不具有免疫力,可以再次被感染。据此可以建立 SEIRS 模型,其微分方程为:

$$\begin{cases} \frac{dS}{dt} = \mu R - \beta SI \\ \frac{dE}{dt} = \beta SI - \sigma E \\ \frac{dI}{dt} = \sigma E - \gamma I \\ \frac{dR}{dt} = \gamma I - \mu R \end{cases} \quad (14)$$

式中: S, E, I, R 分别表示易感者、潜伏者、患病者、康复者在人群中的占比,因此可知 $(S, E, I, R) \in [0, 1]^4$ 。

针对某一地区的某一疾病,根据参数的生物学含义和经验,实验中选取 $D' = [0, 0.5]^4$, 并随机选择参数 $(\beta, \sigma, \gamma, \mu) = (0.8, 0.3, 0.4, 0.3)$ 。当 $\Delta t = 0.2$ 时,生成并构造数据集,驱动 DRNet 和对照网络模型对系统(14)拟合逼近。根据多次试验的损失函数,在 Data-10k 上表现最优的 DRNet 含有 6 个隐藏层、每层 50 个神经元,在 Data-5k 上表现最优的 DRNet 含有 5 个隐藏层、每层 20 个神经元。对于 D' 内任意的初始状态 $(S, E, I, R)_{t=0} = (0.5, 0.5, 0, 0)$, 训练完毕的最优 DRNet 在 $t \leq 20$ 的预测结果如图 5 中圆点所示。由图 5 中结果可见,DRNet 在 3 个数据集上均较好地实现了对 SEIRS 模型的拟合逼近。

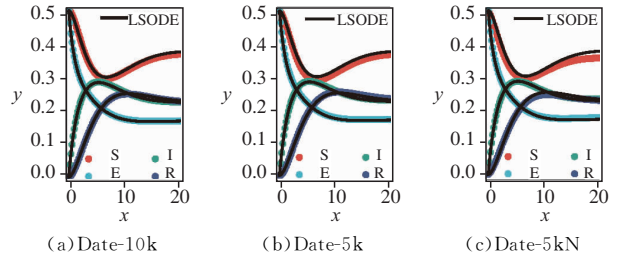


图 5 DRNet 应用于 SEIRS 模型的结果

DRNet 和对照网络模型的评价指标见图 6。从图中的 4 个定量评价指标可知,在 3 个数据集上,DRNet 的表现均优于对照网络模型。

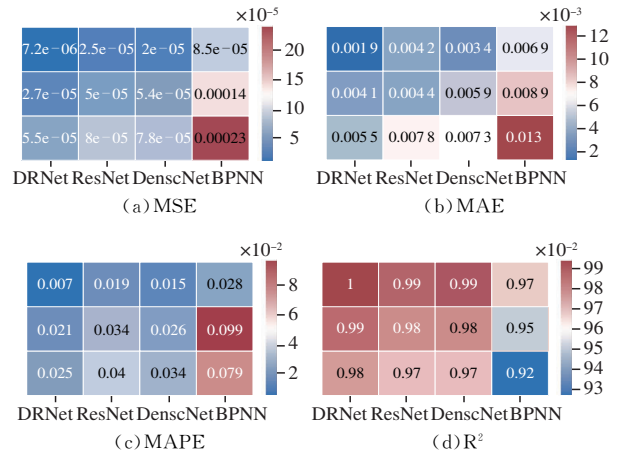


图 6 DRNet 和 ResNet, DenseNet, BPNN 应用于 SEIRS 模型的评价指标

3.3 Logistic-Volterra 模型

20 世纪 20 年代,致力于微分方程应用研究的意大利数学家 Volterra 在研究捕食与被捕食关系时提出了著名的 Volterra 模型。在该模型中考虑系统内捕食者与被捕食者自身的阻滞作用时,为 Volterra 模型加上 Logistic 项,得到 Logistic-Volterra 模型,其微分方程组为:

$$\begin{cases} \frac{dx}{dt} = rx(1 - \frac{x}{m}) - axy \\ \frac{dy}{dt} = -sy(1 + \frac{y}{n}) + bxy \end{cases} \quad (15)$$

选取某 Logistic-Volterra 系统,根据生物学经验随机取定 $(r, s, a, b, m, n) = (1.1, 0.4, 0.2, 0.03, 90, 30)$, $D' = [0, 30] \times [0, 6]$ 。当 $\Delta t = 0.5$ 时,生成并构造数据集,驱动 DRNet 和对照网络模型对系统(15)拟合逼近。根据多次试验的损失函数,在 Data-10k 上表现最优的 DRNet 含有 3 个隐藏层、每层 20 个神经元,在 Data-5k 上表现最优的 DRNet 含有 3 个隐藏层、每层 50 个神经元。对于 D' 内任意初始状态 $\mathbf{X}_0 = (25, 2)$, 训练完毕的最优 DRNet 在 $t \leq 50$ 的预测结果如图 8 所示。图可见,DRNet 在 3 个数据集上均较好地实现了对 Logistic-Volterra 模型的

拟合逼近。

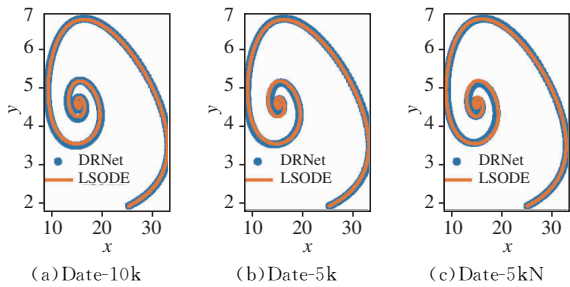


图 7 DRNet 应用于 Logistic-Volterra 模型的结果

DRNet 和对照网络模型的评价指标见图 8。从图中 4 个定量评价指标可知,在 3 个数据集上,DRNet 的预测精度优于对照网络模型。

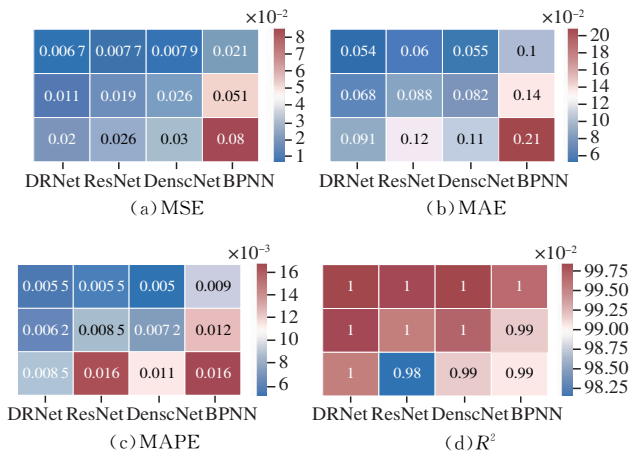


图 8 DRNet 和 ResNet、DenseNet、BPNN 应用于 Logistic-Volterra 模型的评价指标

3.4 结果讨论和参数分析

从图 3、图 5、图 7 可见,经过数据集 Data-5k 和 Data-10k 训练后,DRNet 对系统状态变量的预测值与 LSODE 算子得到的参考值较为契合,表明 DRNet 方法较好地实现了对模型的拟合逼近,具有对状态变量进行准确预测的能力。由图 4、图 6、图 8 中的定量评价指标可见,DRNet 在 2 个数据集上的预测精度均超越了 ResNet、DenseNet 和 BPNN,并且相比于线性的常微分方程系统,DRNet 在非线性的常微分方程系统上的评价指标均比对照网络模型更好,不仅表明 DRNet 在不提高数据规模的条件下,实现了对系统模型的更高精度预测,更表现出 DRNet 方法对自治系统模型逼近和序列预测的高有效性。在抗噪声方面,经过 Data-5kN 训练的 DRNet 虽然性能有所下降,但是依然具有良好的预测精度,均方误差小于对照网络模型,这表明 DRNet 具有较好的抗噪声干扰性能。

在 3 个示例模型中,从数据集 Data-5k 到 Data-10k,最优 DRNet 含有的隐藏层数 h 和每层神经元数量 m 变化不大;且在数据集 Data-10k 的预测精度

要优于其在 Data-5k 的预测精度,这表明可以通过增加数据集规模的方法提升 DRNet 的预测精度。但是对于很多系统而言,获得大量观测数据的成本是巨大的,并且由图 5、图 7、图 9 中的定量评价指标可见,当模型精度已经很高时,大幅增加观测数据集规模对提高预测精度的效果较为有限。

4 结语

本文研究了密集残差网络 DRNet 对自治系统进行模型拟合逼近并实现状态变量预测的问题。在示例模型应用中,DRNet 可实现媲美 LSODE 算子数值解的优异结果,预测精度优于 ResNet、DenseNet 和 BPNN,并且在非线性系统上,DRNet 的定量评价指标明显优于对照网络模型;在面对较小噪声的干扰时,DRNet 的性能没有出现大幅下降,展现出了较好的抗噪声干扰能力。这为利用观测数据研究机理复杂或暂不清楚的系统提供了新的有效方法。

参考文献

- [1] MORENO-SALINAS D, BESADA-PORTAS E, LÓPEZ-OROZCO J A, et al. Symbolic Regression for Marine Vehicles[J]. IFAC-Papers OnLine, 2015, 48(16): 210-216.
- [2] BONGARD J, LIPSON H. Automated Reverse Engineering of Nonlinear Dynamical Systems[J]. Proceedings of the National Academy of Sciences of the United States of America, 2007, 104(24): 9943-9948.
- [3] FUJII K, KAWASAKI T, INABA Y, et al. Prediction and Classification in Equation-Free Collective Motion Dynamics[J]. PLoS Computational Biology, 2018, 14(11): e1006545.
- [4] 崔恒建, 胡涛. 新型冠状病毒肺炎疫情预测预报的非线性回归方法[J]. 中国科学: 数学, 2021, 51(8): 1267-1278.
- [5] LI X, KULANDAIVELU J, ZHANG S, et al. Data-driven Estimation of COVID-19 Community Prevalence through Wastewater-Based Epidemiology[J]. Science of the Total Environment, 2021, 789: 147947.
- [6] QIN T, WU K, XIU D. Data Driven Governing Equations Approximation Using Deep Neural Networks[J]. Journal of Computational Physics, 2019, 395: 620-635.
- [7] SU W H, CHOU C S, XIU D B. Deep Learning of Biological Models from Data: Applications to ODE Models[J]. Bulletin of Mathematical Biology, 2021, 83(3):

- 1-19.
- [8] ZHANG Y L, TIAN Y P, KONG Y, et al. Residual Dense Network for Image Super-Resolution[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City:IEEE, 2018:2472-2481.
- [9] ZHANG Y L, TIAN Y P, KONG Y, et al. Residual Dense Network for Image Restoration [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(7):2480-2495.
- [10] PUROHIT K, MANDAL S, RAJAGOPALAN A N. Mixed-Dense Connection Networks for Image and Video Super-Resolution [J]. Neurocomputing, 2020, 398:360-376.
- [11] NELSON R A, OLSSON M G. The Pendulum-Rich Physics from a Simple System[J]. American Journal of Physics, 1986, 54(2):112-121.
- [12] 张太雷. 几类传染病动力学模型研究[D]. 乌鲁木齐:新疆大学, 2008:1-5.
- [13] BLANCO J M. Relationship between the Logistic Equation and the Lotka-Volterra Models[J]. Ecological Modelling, 1993, 66(3):301-303.
- [14] 马知恩, 周义仓, 李承治. 常微分方程稳定性与稳定性方法[M]. 2 版. 北京:科学出版社, 2015:23-28.
- [15] BAKER M R, PATIL R B. Universal Approximation Theorem for Interval Neural Networks[J]. Reliable Computing, 1998, 4(3):235-239.
- [16] WU K, XIU D. Numerical Aspects for Approximating Governing Equations Using Data[J]. Journal of Computational Physics, 2019, 384:200-221.
- [17] ELDAN R, SHAMIR O. The Power of Depth for Feedforward Neural Networks [J]//Proceedings of Machine Learning Research, Columbia University, 2016(49):907-940.
- [18] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely Connected Convolutional Networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017:2261-2269.
- [19] 李贵锋, 童宁宁, 冯为可, 等. 基于 DenseNet 的机载雷达动目标检测[J]. 空军工程大学学报(自然科学版), 2021, 22(2):83-90.
- [20] 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 (CVPR). Las Vegas; [s. n.], 2016:770-788.
- [21] 杨冰, 徐丹, 张豪远, 等. 基于改进的 Dense Net-BC 对少数民族服饰的识别[J]. 浙江大学学报(理学版), 2021, 48(6):676-683.
- [22] HINDMARSH A C. LSODE. Ordinary Differential Equation System Solver [R]. United States: U. S. Department of Energy Office of Scientific and Technical Information, 1983.

(编辑:徐敏)

(上接第 82 页)

- [11] SUNG F, YANG YX, ZHANG L, et al. Learning to Compare: Relation Network for Few-Shot Learning [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT:IEEE, 2018:1199-1208.
- [12] FINN C, ABBEEL P, LEVINE S. Model-Agnostic meta-Learning for Fast adaptation of Deep Networks [C]//International Conference on Machine Learning. PMLR, 2017:1126-1135.
- [13] YANG N, ZHANG B, DING G, et al. Specific Emitter Identification with Limited Samples: A Model-Agnostic Meta-Learning Approach[J]. IEEE Communications Letters, 2022, 26(2):345-349.
- [14] ZHANG Z, LI Y, GAO M. Few-Shot Learning of Signal Modulation Recognition Based on Attention Relation Network [C]//2020 28th European Signal Processing Conference (EUSIPCO). [S. l.]: IEEE, 2021:1372-1376.
- [15] LIANG Z, MINGLIANG T, WANG L, et al. Automatic Modulation Recognition Based on Adaptive Attention Mechanism and ResNeXt WSL Model [J]. IEEE Communications Letters, 2021, 25(9):2953-2957.
- [16] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, US: IEEE, 2016:770-778.

(编辑:徐敏)