文章编号: 0258-1825(2019)03-0498-07

自动编码器在流场降阶中的应用

叶舒然^{1,2},张 珍^{1,2},宋旭东³,杜特专^{1,2},王一伟^{1,2,*},黄晨光^{1,2},陈耀松³ (1.中国科学院力学研究所流固耦合系统力学重点实验室,北京 100190;

2. 中国科学院大学 工程科学学院,北京 100049; 3. 北京大学 工学院,北京 100871)

摘 要:自动编码器作为一种压缩算法,在数据降维和去噪等方面有着广泛实践,有条件作为一种降阶方法在流场 识别与数据处理方面得到应用。文章中以圆柱绕流为例,首先对圆柱后速度场建立了编码模型,用来对原始数据 进行降维和特征提取,之后将编码后的数据与流场特征量相关联,建立了由流场编码回归圆柱表面压力系数的神 经网络,探索了降维后数据的应用。结果表明,自动编码得到的结果能够承载原始速度场的主要信息,解码后速度 场与原速度场测试均方根误差小于 0.02,压力回归测试均方根误差可小于 0.1。说明自动编码器能够作为一种流 场的特征提取和降阶方法,在未来得到更广泛的应用。

关键词:机器学习;自动编码器;圆柱绕流;流场特征提取;压力预测

中图分类号:TP18; V211.3 文献标识码:A doi: 10.7638/kqdlxxb-2019.0039

Applications of autoencoder in reduced-order modeling of flow field

YE Shuran^{1,2}, ZHANG Zhen^{1,2}, SONG Xudong³, DU Tezhuan^{1,2},

WANG Yiwei^{1,2,*}, HUANG Chenguang^{1,2}, CHEN Yaosong³

 Key Laboratory for Mechanics in Fluid Solid Coupling Systems, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China;

2. School of Engineering Science, University of Chinese Academy of Sciences, Beijing 100049, China;
3. College of Engineering, Peking University, Beijing 100871, China)

Abstract: Known as a compression algorithm, autoencoder is widely used for dimensional reduction and image denoising. It can be used in flow field identification and data processing as a reduced-order method. Moreover, massive labeled flow field data makes it promising to apply machine learning in fluid dynamics. Taking the flow around a cylinder as an example, the autoencoder model of the velocity field behind the cylinder is established to reduce the order and extract features from the original data. This model encodes the 1394 velocity components into 32dimensional data, and trains them through self-supervised learning. For well-trained autoencoder models, the basis for evaluating them is to discuss whether the original flow field could be reconstructed. The application of the low dimensional encoded data is explored by correlating it with the flow field sensitive outputs, and a neural network for the regression of surface pressure of cylinder based on encoded data is established. It is verified that the result of the autoencoder has inherited the main information in the original velocity field. The root mean square error of the decoded velocity field compared with the original field is less than 0.02, and the root mean square error of the pressure coefficient regression network can be less than 0.1. The above results indicate that the autoencoder can be used in the future as a feature extraction and order reduction method of flow field.

收稿日期:2019-03-29; 修订日期:2019-04-26

基金项目:国家重点研发计划(2016YFC0300600,2016YFC0301601)

作者简介:叶舒然(1995-),女,安徽合肥人,硕士研究生,研究方向:水动力学. E-mail:yeshuran@imech.ac.cn

通信作者: 王一伟*(1983-), 男, 研究员, 研究方向: 水动力学. E-mail: wangyw@imech.ac.cn

引用格式:叶舒然,张珍,宋旭东,等.自动编码器在流场降阶中的应用[J].空气动力学学报,2019,37(3):498-504.

doi: 10.7638/kqdlxxb-2019.0039 YE S R, ZHANG Z, SONG X D, et al. Applications of autoencoder in reduced-order modeling of flow field[J]. Acta Aerodynamica Sinica, 2019, 37(3): 498-504.

Keywords: machine learning; autoencoder; flow around a cylinder; flow field feature extraction; pressure prediction

0 引 言

机器学习作为计算机科学领域中一个备受关注的研究方向,其核心主要是使用算法解析数据,并利用计算机的大信息处理能力从原始数据中提取模式。近来,机器学习在许多领域得到蓬勃发展,例如在自然语言处理和各类图像识别等方面已经得到广泛应用^[1]。机器学习在流体力学中的应用也开始逐渐浮现,Tracey等^[2]利用神经网络拟合 Spalart-Allmaras 湍流模型并证明了其能够优化计算流体力学中得到的结果;Ling等^[3]提出使用嵌入伽利略不变性的多 层网络预测雷诺应力张量,并取得了更为精确的结 果。至此,机器学习在流体力学中的初步尝试主要是 通过前馈神经网络构建湍流模型。

由于流场的复杂性,除了关注于传统前馈网络学 习算法,利用机器学习中的各种方法来探索流场识 别、提取、降阶也是一个非常重要的方向。以卷积神 经网络为代表的深度网络在流体力学中有了一些新 的尝试。Storfer 等^[4]使用卷积神经网络成功识别了 流动特征,并且识别出的特征能够与相似特征进行区 分。Jin 等^[5]构建了一个捕捉绕流圆柱表面压力系数 时空信息的卷积神经网络,能够成功捕捉压力脉动特 征并预测速度场信息。对于流场建模,除了直接寻找 数据间的映射关系,基于无监督思想下的寻找数据之 间的特征提取技术也是一种建模与降阶技术[6]。例 如 Kaiser 等^[7]提出了一种无监督学习下用以识别流 场物理机制的 CROM 方法,这种方法在三维钝体湍 流尾迹和空间演化的不可压混合层的速度场都进行 了成功的应用。考虑到自动编码器作为一种典型的 数据压缩与降维方法,在图像和自然语言处理等方面 有很好的实践,因此探索自动编码器在流场降阶中的 应用。将自动编码器引入流场,建立流场速度场的编 码和解码模型,能够将原始高维数据进行压缩,并且 能够对流场实现一种通用的降阶方法。

本文以圆柱浇流流场的速度场为输入,建立了自 动编码器模型,对全场的速度场分量进行编码和数据 降维,最后与流场特征量构建的网络进行了误差分 析,分析了自动编码器的数据降维和特征提取的合理 性。该方法在一定范围内的速度入口和雷诺数条件 下,能够为流场数据的表示提供合适的编码函数。并 且,结合深度学习技术,能够得到相关的流场特征识 别与提取结果。

1 自动编码器原理与设计

自动编码器是一种神经网络,这种网络的特点是 在经过训练后能够将输入复制到输出^[8]。作为无监 督的一种网络模型,自动编码器能够从输入数据中学 习到隐含的数据特征,并且由这些特征重构出原始数 据。因此,自动编码器在数据降维和特征学习等方面 有着广泛的应用。

1.1 前馈网络结构

自动编码器可以被看作是前馈网络的一个特例^[9]。因此,首先建立一个全连接的前馈网络,其网络结构如图1(a)所示,分为输入层、隐层、输出层,上一层的任何神经元与下一层的每个神经元都有连接^[10]。神经元模型的概念来自于生物神经中,通过电位变化传递信息^[11]。对于神经网络模型中的神经 元模型,其权重函数和偏置向量,用方程表示为:

 $f(x;w;b) = x^{T}w + b$ (1) 式中,w 为权重系数,b 为偏置向量。

同时,为了使每个神经元模型表现为非线性,在 上述变换后增加一个非线性激活函数,如 ReLU: $\sigma(z) = \max \{0, z\},$ 或 sigmoid 函数: $\sigma(z) = \frac{1}{1+e^{-z}}$ ^[12]。 不同的连接方式可以构成不同的网络结构,也称为神 经网络设计的架构,包括网络中的单元数量以及这些 单元之间的连接方式。神经元被组成层的单元组的 链式结构,层与层之间通过全连接相连。

神经网络将输入层数据通过映射转换到隐层,再



逐层转化到输出层。网络的输出和真实值之间的差 异用损失函数衡量^[13],对于常见的回归问题,一般取 为网络预测值和参考值之间的均方根误差:

$$\delta_{\text{RMSE}} = \sqrt{\sum_{i=1}^{n} (y_i - \tilde{y}_i)^2 \cdot \frac{1}{n}}$$
(2)

其中 y_i 为参考值, y_i 为网络预测值, n 为样本数。 网络训练是通过对损失函数的梯度下降和误差的反 向传播,使网络在训练过程中逐步拟合训练数据。通 常也可以采用 Dropout 技术,提高网络的泛化能力, 防止网络对训练数据过拟合^[14]。

1.2 自动编码器设计

500

前馈神经网络,如果其只有一个隐藏层h,则该 网络可以看作为从输入到隐层的映射函数h = f(x)和隐层到输出的映射函数y = g(h)。由于自动编码 器的特点是能够将输入复制到输出,所以在网络中将 输出设计成与输入相同,并将隐层的节点数设置成比 输入输出的节点数少,通过训练调整权重系数,得到 每一层的参数,便可以获得原始数据在小维度上的表 示。对于该网络而言是一个编码和解码的过程,即将 原始数据进行压缩,再从压缩的数据中还原。

图 2 是自动编码器的网络结构,在该网络中输出 与输入相同,即 y=x。自动编码器的训练方式和一 般前馈神经网络完全相同,即使用小批量的梯度下降 算法,使输出 y 逐渐贴近输入 x^[15]。网络训练完成 后,能够得到一个输入到隐层的编码模型和一个隐层 到输出的解码模型,隐层的节点信息即为最后的编码 结果。考虑到编码的稀疏性,在自动编码器的基础上 加上稀疏性约束,编码结果中只有一小部分神经元被 激活,大部分节点为 0^[16]。



2 圆柱绕流自动编码模型建立

2.1 圆柱绕流算例

选取圆柱绕流的案例,构建流场并获取流动数据^[17]。流场几何设置如图 3(a)。考虑圆柱后 41×17

的速度场作为输入。由于是二维算例,对于每个点都 有 U={u,v}两个速度分量。流场的速度入口,考虑 了带有扰动的抛物速度入口条件,其中3个速度入口 作为训练,第四个速度入口作为测试。每个速度入口 考虑雷诺数从 200 到 2000 的 10 组雷诺数,所有工况 下均每隔 0.005 s 时间步长输出一组数据,共取 300 组数据。速度入口条件为基本抛物线速度入口乘以 扰动因子 P 再归一化,基本抛物线速度入口为:

$$u_{s}(0,k) = \frac{4u_{m}k(H-k)}{H^{2}}, v = 0, \qquad (3)$$

其中 $u_m = 1.5$ 。扰动因子 P 见表 1。

表1 圆柱绕流流场不同速度入口条件

Table 1 Different inlet velocity conditions of the flow around the cylinder

速度入口	数据集	扰动因子 P
速度入口1	训练集	$1 + \sin \left(\frac{2\pi}{0.41}y\right)^2$
速度入口 2	训练集	$1 + \sin \left(\frac{3\pi}{0.41}y\right)^2$
速度入口 3	训练集	$1 + \cos(\frac{6\pi}{0.41}y)^2$
速度入口 4	测试集	$1 + \sin \left(\frac{6\pi}{0.41}y\right)^2$



(a) 几何尺寸和边界条件



501

网络分为自动编码器和验证网络两部分,自动编码器考虑将 41×17×2 个速度分量通过自动编码,压 缩为 32 个数据。验证网络将用编码后的数据与流场 特征典型量相联系,讨论压缩处理后的数据是否抓住 了原始数据中的主要特征。

2.2 自动编码器结构

对于圆柱绕流流场,采取上述的速度场进行描述,考虑到是二维流场,则在 41×17 个空间位置上共 有 1394 个速度分量,对于其余的研究而言速度信息 的维度过大,因此,寻找一种将 1394 维的流场速度数 据进行压缩的方法十分重要。自动编码器结构能够 将数据进行压缩,并且能够利用网络将数据还原。考 虑到自动编码器的这个优势,使用自动编码器将流场 速度分量压缩为只有 32 个值。因此对于网络,输入 和输出选择各为 1394 个节点,中间隐层选择为 32 个 节点,网络中通过选取不同的激活函数来添加约束, 使得重构后的数据与原始数据非常接近但又有一些 不同,以此强制模型考虑输入数据中哪些量需要被优 先复制,达到增强网络性能的目的。

2.3 编码器验证与应用方法

对于自动编码器得到的缩减结果,一个很好的验 证方式是,将经过解码后的数据与原始数据比较,得 到两者间的均方根误差(Root Mean Square Error, RMSE)。由于流场信息的复杂性,想要探究缩减后 的结果对于原始数据是否具有代表性,另一个验证方 法是建立编码后数据与原始流场的对应关系。如果 编码后的数据能够代替原始范围内的速度场信息,那 么可以建立函数关系 $C_p = F(U_{encode})$,其中 C_p 是用 入口平均速度无量纲化的压力系数,Uencode 为编码后 的低维流场。基于这种考虑,建立了全连接的神经网 络,输入层为编码后的流场速度,输出层为圆柱表面 压力系数。如图 3(c)所示,圆柱表面压力系数选取 为圆柱一圈 32 个离散点的值进行表示。由于训练集 和测试集选取了不同的速度入口,为了防止入口速度 变化对圆柱前缘处压力系数影响较大,在32个点中 舍弃圆柱了迎着来流的6个点。圆柱表面压力采样 点排列顺序是以迎来流最前端为第一个点顺时针等 间距排列,因此在序号上体现为舍弃前后端的三个, 并对剩下的 26 个点按照相同的顺序重新编号为 1~ 26。这些点仍能覆盖圆柱表面大部分压力情况,并且 不会使得速度入口所产生的误差对整体误差起到了 主导作用。验证网络选取了单隐层的 128 个节点的 验证网络1,和隐层的数量为2层、每层30个节点的 验证网络 2,并加了 dropout 层以防止过拟合。验证 网络结构见表 2。

表 2 验证网络的网络结构

Table 2 Network layers of the verification network

验证网络1		验证网络2	
输入层	32	输入层	32
全连接层1	128	全连接层1	30
随机失活层	0.2	随机失活层	0.2
输出层	26	全连接层 2	30
		随机失活层	0.2
		输出层	26

3 圆柱绕流自动编码器的应用

输入训练数据完成网络训练后,需要应用测试数 据测试网络的可靠性。本文的自动编码器在流场中 的应用主要分为两个方面,一方面关注了自动编码器 编码解码过程中的精度,是否能生成与原始流场精度 差异较小的重构流场,另一方面也考虑到编码过程是 否能够对原始流场进行成功的压缩降阶,并通过关联 流场中的某些敏感关注量考察编码器能否代替原始 流场。具体流程见图 4。



Fig.4 Flow chart of the application of the autoencoder

3.1 自动编码器的表现

上述圆柱绕流案例有 10 组雷诺数,每组雷诺数 每隔 0.005s 取一组数据,共计算了 300 步数据。为 了防止训练集对于速度入口过拟合,共考虑了 4 个不 同的速度入口,3 组进行训练,1 组进行测试,即训练 样本数为 9000,测试样本数为 3000。

自动编码器的结构采用的是 2.2 节中所描述的 网络结构。图 5 可以看出,随着训练步数的增加,训 练过程中的均方根误差即损失函数在逐渐下降,经过 100 步的迭代误差已经降到 1.2×10⁻²,并有继续下降 的趋势。

网络训练完成后,将测试数据的速度场输入网络 模型,即可得到测试数据在自动编码后的系数,其均 方根误差 δ_{RMSE} 为1.45×10⁻²。可以认为测试数据在 经过编码和解码之后能得到与原始数据误差较小的 结果。随机选取 *Re* 数为 700 的一个速度场进行研 究,从速度云图上看,其中图 6(a)为原始速度场的速 度 *u* 分量的速度云图,图 6(b)为重构速度场的速度 u分量的速度云图,可以看出该编码解码过程误差较小,说明编码后的数据能够包含大部分流场信息。

同时,我们将编码维数增加一倍,使用自动编码 器将流场速度分量压缩为 64 个值,并使用 64 维数据 重构流场,如图 6(c)。可以看出编码维数的增加,能 够使得重构出的流场更好地呈现原始流场的流动特 征,例如在核心区外围的情况能够将流场信息把握得 更好。





(a) 原始流场



(b) 重构流场 1







3.2 基于压力回归的网络验证

自动编码器训练和测试完成后,进一步利用前述 的压力回归方法对编码器表现进行验证。对于验证 网络,训练集和测试集的数据来源依旧是上述雷诺数 和速度入口组合下的工况。对于每组数据,输入由原 始的 1394 个速度分量缩减为编码后的 32 个组合值, 输出则考虑到需要对原始流场信息进行表征,因此选 取了圆柱表面 26 个点的压力系数。分别用 2.3 节所 述的两种网络结构进行了验证,其中编码后数据在验 证网络上的表现如表 3 所示。

表 3 训练集和测试集分别在验证网络上的均方根误差

 Table 3
 RMSE for the train set and the test set on

verification network

	验证网络1	验证网络 2
训练集	2.48×10^{-2}	$7.14 imes 10^{-2}$
测试集	0.129	0.111

3.3 与随机取点对比

为了验证自动编码的效率,将流场的41×17×2 个速度分量中随机选取32个点作为对照组,其中随 机选取的过程保证所取点在训练数据和测试数据中 代表的意义相同,即在固定空间位置上的相同速度 分量。

通过对比 2.3 节所建立网络与随机取点,分别得 到了单层编码结果和随机取点在验证网络上的均 方根误差。表 4 中对比了两种网络上测试集的误 差。图 7 选取了某时刻预测的压力系数与实际值的 对比。自动编码器的残差约为随机取点网络的一半, 说明了自动编码器的压缩算法能够更有效表达原流 场的信息。

表 4 单层编码和随机取点在验证网络上的均方根误差 Table 4 RMSE for the encoded result and the random-sampling on verification network

	验证网络1	验证网络 2
单层编码	0.129	0.111
随机取点	0.246	0.208



图 7 单层编码与随机取点在验证网络上预测结果与 CFD 参考值结果中圆柱压力系数分布的比较

Fig.7 Comparison of pressure coefficient distributions on the cylinder between the CFD result, the model predictions of the encode result and the random sampling

4 编码层层数影响分析

4.1 网络结构

前文中采用的是基本的单层编码器和解码器,但 实际上对于复杂问题,深度的编码器和解码器也具有 一定优势,能够更好地体现深层网络的可控性,在拟合 数据过程中较大地降低训练的计算成本和所需的训练 量。例如,Hinton 等^[18]提出在实验中,深度自编码器 能够比相应浅层的编码器拥有更好的压缩效率。

对于速度场信息压缩的问题,为了探究自动编码 器网络对编码效率的影响,提出了一个隐层数为5的 网络,新的编码器网络结构如表5。在该网络中,编 码器和解码器均为3层网络。

表 5 自动编码器的网络结构

 Table 5
 Network Layers of the autoencoder

网络层	节点数	激活函数
输入层	1394	
全连接层1	256	ReLU
全连接层 2	64	ReLU
全连接层 3	32	ReLU
全连接层 4	64	ReLU
全连接层 5	256	Tanh
输出层	1394	

4.2 结果对比与讨论

训练深度编码模型,其本质是训练具有更多中间 层的网络。对于流场速度场数据,网络的输入是 1394 维特征,与单层编码不同的是,深度编码网络中 三个隐含层维度分别为 256、64 和 32,即通过逐步降 低输入数据的维度,得到最终编码后的 32 维的低维 数据。表 6 对比了单层编码和深层编码在编码网络 (AE-Net)和验证网络(Verif-Net)上的均方根误差, 可以看到,在最终维度相同的情况下,深层编码与单 层编码相比,重构出的数据误差更小,且能够更好地 捕捉数据的特征,压力回归测试均方根误差小于 0.1。 但由于目前算例相对简单,深度编码器体现出的优势 并不十分明显。

表 6 单层编码和深层编码网络的均方根误差

 Table 6
 RMSE for the single layer encoded result

 and the multi-layer encoded result

均方根误差	编码网络	编码网络	验证网络1	验证网络 2
	训练误差	测试误差	测试误差	测试误差
单层编码	$1.20 imes 10^{-2}$	$1.45 imes 10^{-2}$	0.129	0.111
深层编码	$6.86 imes 10^{-4}$	1.01×10^{-2}	0.107	0.096

5 结 论

本文提出了一种将自动编码器应用在流场中的

方法。该方法通过将高维的原始速度场编码为少量 点的低维数据,并对数据解码,能够对原始流场进行 重构。对非均匀来流圆柱绕流算例建立的典型单层 自动编码器模型,重构流场与原始流场速度均方根误 差小于 0.02。

本文证实了通过自动编码器将原始高维流场进 行降阶和数据压缩,可以将压缩后数据与原流场中的 敏感输出值相关联。对此基于圆柱绕流流场自动编 码器的应用,建立了利用降维后编码数据来回归圆 柱表面压力分布的神经网络测试。结果表明编码 数据能够较好地回归到圆柱的表面压力系数,且回归 精度明显高于利用同等数量随机测点进行回归的对 照模型。

最后考虑到更复杂流场的编码需求,讨论了编码 层层深对于编码效果的影响。典型的不同结构编码 器对比结果表明,五层结构的深层编码器比浅层编码 器在编码网络上的误差更低,编码精度更高;同时在 与流场敏感输出值相关联时网络预测误差能够达到 0.1,对原始数据的代表性也更好。

未来,本文所建立的基于机器学习的自动编码器 方法有条件应用于更复杂的流场结构的识别与降阶, 应用的范围与方式值得进一步讨论研究。此外,如将 自动编码器与潜变量模型理论结合,则可进一步应用 于流场生成模型的构造,值得进一步探索。

参考文 献:

- [1] LECUN Y, BENGIO Y, HINTON G. Deep learning. [J]. Nature, 2015, 521: 436-444.
- [2] TRANCY B D, DURAISAMY K, ALONSO J J. A machine learning strategy to assist turbulence model development[R]. AIAA 2015-1287, 2015.
- [3] LING J, KURZAWSKI A, TEMPLETON J. Reynolds averaged turbulence modelling using deep neural networks with embedded invariance [J]. Journal of Fluid Mechanics, 2016, 807:155-166.
- [4] STROFER, CARLOS M, WU J, et al. Data-driven, physicsbased feature extraction from fluid flow fields [J]. Communications in Computational Physics, 2018, 25(3): 625-650.
- [5] JIN X, CHENG P, CHEN W L, et al. Prediction model of velocity field around circular cylinder over various Reynolds numbers by fusion convolutional neural networks based on pressure on the cylinder[J]. Physics of Fluids, 2018, 30(4): 047105.
- [6] FARAHAT A K, GHODSI A, KAMEL M S. Efficient greedy feature selection for unsupervised learning[J]. Knowledge and Information Systems, 2013, 35(2):285-310.
- [7] KAISER E, NOACK B R, CORDIER L, et al. Cluster-based reduced-order modelling of a mixing layer[J]. Journal of Fluid Mechanics, 2014, 754:365-414.

- [9] LIOU C Y, HUANG J C, YANG W C. Autoencoder for words [J]. Neurocomputering, 2014, 139:84-96.
- [10] ILIADIS M, SPINOULAS L, KATSAGGELOS A K. Deep fully-connected networks for video compressive sensing [J]. Digital Signal Processing, 2017;S1051200417302130.
- [11] ZHANG Z, SONG X D, YE S R, et al. Application of deep learning method to Reynolds stress models of channel flow based on reduced-order modeling of DNS data[J]. Journal of Hydrodynamics, 2018, 31(11): 58-65.
- [12] NAIR V, HINTON G E. Rectified linear units improve restricted Boltzmann machines [C]//Proceedings of the 27th International Conference on Machine Learning. Haifa, ILMS: 2010:807-814.
- [13] CHAI T, DRAXLER R R. Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature[J]. Geoscientific Model Development,

2014, (7):1247-1250.

- [14] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Image net classification with deep convolutional neural networks [C]// Neural Information Processing Systems-Volume 1 (NIPS'12), 2012: 1097-1105.
- [15] 陈海,钱炜祺,何磊.基于深度学习的翼型气动系数预测[J]. 空气动力学学报,2018,36(2):294-299.
 CHEN H, QIAN W Q, HE L. Aerodynamic coefficient prediction of airfoils based on deep learning [J]. Acta Aerodynamica Sinica, 2018, 36(2):294-299. (in Chinese)
- [16] SUN W, SHAO S, ZHAO R, et al. A sparse auto-encoderbased deep neural network approach for induction motor faults classification[J]. Measurement, 2016, 89(ISFA):171-178.
- [17] SCHAFER M, TUREK S, DURST F, et al. Benchmark computations of laminar flow around a cylinder[J]. Notes on Numerical Fluid Mechanics, 1996, 48:547-566.
- [18] HINTON G E, SALAKHUUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313: 504-507.