

DOI: 10.11992/tis.201905044

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20190909.1705.006.html>

## 图像情境下的数字序列逻辑学习

梁慧<sup>1,3</sup>, 曹峰<sup>1,3</sup>, 钱宇华<sup>1,2,3</sup>, 郭倩<sup>1,3</sup>, 梁新彦<sup>1,3</sup>

(1. 山西大学 大数据科学与产业研究院, 山西 太原 030006; 2. 山西大学 计算智能与中文信息处理教育部重点实验室, 山西 太原 030006; 3. 山西大学 计算机与信息技术学院, 山西 太原 030006)

**摘要:** 针对未知的数字和规则的模式构建问题, 本文提供了一种从图像角度解决数字序列逻辑学习问题的方法。该方法是在计算机不知道图像间关系和图像内包含的内容的意义的前提下, 让计算机自主地学习出其中包含的内在逻辑模式, 从而进行数字序列的预测。本文构建了 4 个大型数据集: Linear 序列、Multiplication 序列、Fio 序列和 Nested 序列, 然后使用几种代表性的深度神经网络来完成数字序列逻辑学习任务, 并对实验结果加以分析比较, 事实证明, 本文所提出的方法在一定程度上可以解决未知的数字和规则的模式构建问题, 这为一系列未知逻辑模式构建任务提供了一种可能性。

**关键词:** 人工智能; 逻辑推理; 逻辑学习; 深度学习; 数字序列; 图像处理; 神经网络; 模式构建

**中图分类号:** TP181 **文献标志码:** A **文章编号:** 1673-4785(2019)06-1189-10

中文引用格式: 梁慧, 曹峰, 钱宇华, 等. 图像情境下的数字序列逻辑学习[J]. 智能系统学报, 2019, 14(6): 1189-1198.

英文引用格式: LIANG Hui, CAO Feng, QIAN Yuhua, et al. Number sequence logic learning in image context[J]. CAAI transactions on intelligent systems, 2019, 14(6): 1189-1198.

## Number sequence logic learning in image context

LIANG Hui<sup>1,3</sup>, CAO Feng<sup>1,3</sup>, QIAN Yuhua<sup>1,2,3</sup>, GUO Qian<sup>1,3</sup>, LIANG Xinyan<sup>1,3</sup>

(1. Research Institute of Big Data Science and Industry, Shanxi University, Taiyuan 030006, China; 2. Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan 030006, China; 3. School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China)

**Abstract:** To solve the problem of pattern construction of unknown numbers and rules, in this paper, we provide a method to solve the problems of number sequence logic learning from the image perspective. The method allows the computer to automatically learn the inherent logic pattern without prior knowledge of the meaning of the image content or of the relationship between images so as to predict the number sequence. Four large datasets were constructed: linear sequences, multiplication sequences, fio sequences, and nested sequences, and then several representative deep neural networks were used to complete the number sequence logic learning task. By analyzing the experimental results, the method was found capable of solving the problem of pattern construction for unknown numbers and rules to a certain extent, which will provide a potential solution for a series of unknown logic pattern construction tasks.

**Keywords:** artificial intelligence; logical reasoning; logical learning; deep learning; number sequences; image processing; neural network; pattern construction

收稿日期: 2019-04-15. 网络出版日期: 2019-09-10.

基金项目: 国家自然科学基金项目 (61672332, 61432011, U1435212, 61872226); 山西省海外归国人员研究项目 (2017023); 山西省自然科学基金计划资助项目 (201701D121052).

通信作者: 钱宇华. E-mail: [jinchengqyh@126.com](mailto:jinchengqyh@126.com).

数字序列逻辑学习的发展可以说是经历了漫长的岁月, 其可以追溯到人类早期历史。例如, 古时满月的预测<sup>[1]</sup>, 并且其仍然是当今研究的活跃领域, 在股市中, 我们经常会听到有神奇数字时间之窗这一说法, 其用到的波浪理论的数字基

基础就是一系列的数列,此数列称为 fibonacci 序列。大自然中还有很多与 fibonacci 序列有关的奇妙现象,如蜘蛛网、水流的旋涡、蜗牛壳的螺旋以及星系内星球的分布等。

数字序列逻辑学习问题是归纳推理和模式发现的经典问题之一<sup>[2-3]</sup>,这些问题在哲学和数学领域中已经研究了数千年,在心理学和计算机科学领域中也已经研究了数十年<sup>[4]</sup>。近些年来随着机器学习和深度学习的不断发展<sup>[5-8]</sup>,序列预测问题更是吸引了大家的广泛研究。数字序列预测问题在各个领域的应用也是相当的广泛。其可以编码渐进矩阵问题<sup>[9]</sup>,该问题是用二维矩阵表示的,科学预测问题包括股票的预测、智商测试等,以及许多归纳推理问题<sup>[10-12]</sup>。早在 1963 年,Simon 等<sup>[13]</sup>就提出了针对一系列字母序列来预测的模式描述程序,并解释了人类如何从记忆中的概念或规则产生连续模式。Sanghi 等<sup>[14]</sup>提出了一种用于智力测试的程序,该程序还解决了各种数字序列预测问题。且针对数字序列预测问题的解决已有很多方法,例如:反统一算法成功应用于交替序列和斐波纳契序列。Siebers 等<sup>[15]</sup>提出了一种半分析方法用以解决自然数序列归纳问题,这实则是一种典型的智力测试任务,该方法通过术语结构的启发式枚举来猜测给定数字序列的术语结构,最后利用数字序列评估系统来评估该方法,最终结果精度达到 93.2%。然而,其缺点是枚举和搜索受到一种看似合理的偏见的严格限制。Strannegard 等<sup>[16]</sup>提出了 IQ 测试中数字序列预测问题的计算方法。该计算方法是通过开发 ASolver 来实现的,ASolver 是一种基于有限工作记忆思想的拟人化认知系统,其利用了人类推理的模型。结果表明,该算法的性能优于 Maple 和 Wolfram-Alpha 等数学工具。Hofmann 等<sup>[17]</sup>证明了归纳程序系统 IGOR2 可以解决数字序列预测问题,其不同于上述专门用于解决数字序列域中问题的系统,IGOR2 可以解决不同问题解决域中的问题。IGOR2 是一个从小组输入/输出示例中学习功能程序的系统。但不足之处在于其无法很好地处理负数和交替序列。Ragni 等<sup>[18-19]</sup>提出了一种基于人工神经网络 (ANNs) 的动态学习方法来解决数字序列的预测问题。其总体结果与人类水平相当,但误差分布差异很大,且无法解决 Nested 序列问题(见 2.1)。

上述方法都是传统的方法,它是基于已知的数字和规则来构建模式,然后预测数字。然而,

传统数字序列逻辑学习也存在一定的不足,它很难解决未知数字和规则的模式构建的数字序列预测问题。例如,每月的天数序列、Nested 序列、交替序列、带负数的序列等。举一个更为具体的例子:1) 2,4,6,8,10,12,..., 2) 1,2,3,5,4,4,2,..., 对于第 1 个序列,可以明显地观察到其是一个加 2 的规律,这样就可以构建一个满足该规律的模式,传统数字序列技术可以很容易地解决此类预测问题。对于第 2 个序列,难以看出其中的规律,且很难用传统方法得到一种满足此序列的模式来预测该序列,实际上该序列是书写汉字如一二三四等需要的笔画数,故而下一个预测结果应为 2。

针对传统数字序列逻辑学习的不足,本文提出了图像数字序列逻辑学习任务,以从另一个角度来解决数字序列预测问题,其可以很好地解决传统数字序列难以解决的未知数字和规则的模式构建的序列预测问题。在传统数字序列逻辑学习中,数字的含义及其规则都被事先定义。而本文提出的图像数字序列逻辑学习是在不知道图像间关系和图像内包含的内容的意义的前提下,仍可以自动学习出其包含的内在逻辑模式。在本文中,图像序列中的所有图像都是 15×85 像素的黑白图像,且每张图像包含一个多位数。然后将图像传递给计算机,且并不告诉计算机图像中包含的内容是什么,让计算机自动学习图像之间的内在逻辑模式,并且预测下一张图像中的内容。由此看来,图像数字序列逻辑学习不仅可以解决数字序列预测问题,更是为解决一系列未知逻辑模式构建任务提供了一种可能。

## 1 研究方法

本节主要介绍了图像数字序列逻辑学习的基本任务,并详细描述了 4 种有代表性的深度神经网络的网络结构 (MLP、LSTM、CNN\_MLP、ResNet)。

### 1.1 基本任务

图像数字序列逻辑学习是从另一个角度来解决数字序列预测问题,其与传统数字序列解决方案不同。传统数字序列技术是在已知数字和规则的前提下构建模式。而本文的方法是在不知道图像间关系和图像内包含的内容的意义下来预测序列,然后利用深度神经网络来完成数字序列逻辑学习任务,该任务的输入是几张相关的图像序列,其输出仍然是图像,展示的是下一张图像序列预测的结果。

首先将图像序列传送给计算机, 让计算机在不知道图像间关系和图像内包含的内容的意义的前提下自动学习出其内在逻辑模式, 从而进行数字序列的预测。所有这些图像都是黑白图像, 右对齐, 大小归一化 ( $15 \times 85$ ), 每张图像包含一个多位数, 其可以是正数或负数, 每张图像中的最大位数设置为 10。如图 1 所示, 其中图 1(a) 的位数设置为 9 位, 图 1(b) 位数设置为 10 位, 图 1(c) 位数设置为 11 位, 原本数字为 32 282 246 720, 由于超出位数的设定, 图像中最左位的数字无法完全显示, 图 1(d) 为负数情形, 位数设置为 10 位 (除符号位)。然后本文使用几种代表性的深度神经网络来完成数字序列逻辑学习任务, 并加以比较分析。给出网络结构的详细描述。最后, 为了测试本文的网络性能, 本文以某种逻辑关系随机生成一批测试样本, 每个样本包括  $k$  张图像, 将前

$k-1$  张图像传入模型中, 测试生成的图像是否正确。



图 1 图像展示  
Fig. 1 Image display

## 1.2 网络结构

本文在构造的数据集上比较了几种有代表性的深度神经网络的性能 (MLP、LSTM、CNN、MLP、ResNet)。数据集的构造详见 2.1。全部模型均使用均方误差 (MSE) 损失作为优化函数, ADAM<sup>[20]</sup> 作为优化器, 衰减率参数设置为  $\beta_1=0.9$ ,  $\beta_2=0.999$ 。每个网络的详细结构和超参数设置如图 2。

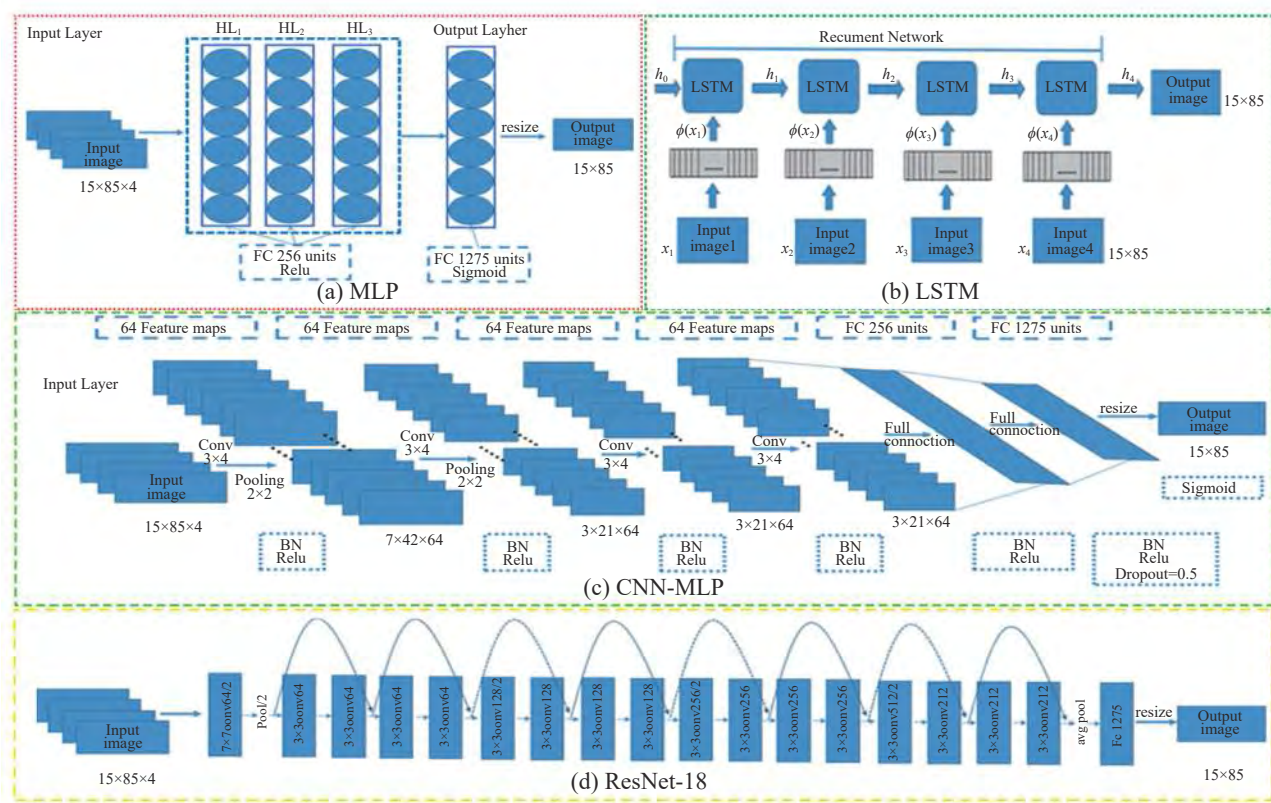


图 2 4 种神经网络结构

**Fig. 2 Four neural network structures.**

多层感知机 (MLP) 在本文中, 使用一个 4 层的多层感知机网络来构建模型, 该模型包含输入层、隐藏层和输出层。采用一个 3 层架构的隐藏层, 每个隐藏层将线性整流函数 (ReLU)<sup>[21]</sup> 作为激活函数。使用 Sigmoid 函数作为网络的输出层。网络结构如图 2(a) 所示。

**长短期记忆网络 (LSTM)** 本文采用标准的 LSTM 网络结构,它是一种特殊的 RNN,主要是为了解决长序列训练过程中的梯度消失等问题<sup>[22]</sup>。而且由于 LSTM 是按顺序地接收输入数据,这非常适合于本文的序列预测问题。本文将图像特征向量  $\phi(x_i)$  按顺序地传入 LSTM 中以编码隐藏状



态  $h_i$ , 其输出  $h_i$  依赖于以前的状态  $h_{i-1}$ , 当前的观察  $\phi(x_i)$  以及上一时刻的单元状态  $c_{i-1}$ 。其单元状态和隐藏状态公式如下:

$$c_i(\phi(x_i), c_{i-1}) = f \cdot c_{i-1} + i \cdot \bar{c}_i \quad (1)$$

$$h_i(\phi(x_i), h_{i-1}) = o \cdot \tanh(c_i) \quad (2)$$

式中:  $c_i$  是  $t$  时刻的单元状态;  $f$ 、 $i$  和  $o$  分别表示遗忘门、输入门和输出门;  $\bar{c}_i$  表示当前输入的单元状态。结构如图 2(b) 所示。

**卷积-全连接神经网络 (CNN-MLP)** 本文使用一个标准的 4 层卷积神经网络<sup>[23-24]</sup>, 并将 ReLU 非线性函数作为卷积层的激活函数, 在卷积输出层后连接一个 2 层的全连接层 (MLP)。为防止过拟合在第一个全连接层上使用丢失率为 0.5 的 dropout<sup>[25]</sup> 技术。结构及其超参数设置如图 2(c) 所示。

**深度残差网络 (ResNet)** 深度残差网络有很多不同的结构<sup>[26]</sup>, 包括 ResNet18、34、50、101、152 等, 在本文实验中, 这几种网络的实验效果相差无几, 所以在这里仅展示了 ResNet18 的实验结果。其结构如图 2(d) 所示。

本文的具体步骤如下: 在训练阶段, 传递给计算机  $N$  张图像, 每张图像包含一个多位数, 如图 3 所示 (假设  $N=4$ ), 而且并不告知计算机每张图像中包含的内容是什么, 然后使用以上所提及的深度神经网络来完成数字序列逻辑学习任务, 以便计算机可以从前 3 张图像间的逻辑模式自动学习第 4 张图像。在测试阶段, 以某种逻辑关系随机生成一批测试样本, 每个样本包括 4 张图像, 将前 3 张图像放入模型中, 以查看是否可以生成正确的预测图像。

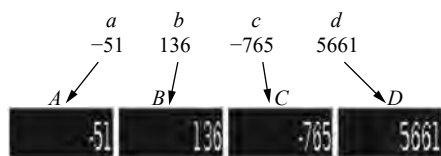


图 3 计算机不知道给定图像中所包含的内容是什么  
Fig. 3 The computer doesn't know what the content in the given image is

对于数据分析, 本文系统地改变学习速率、批量大小、隐藏节点的数量和训练迭代等。这些变化应该允许将不同的神经网络模型与经验结果进行比较。此外, 本文将在所构造的数据集上对比分析不同的深度神经网络模型的性能, 进一步验证本文所提出的图像数字序列逻辑学习对未知

数字和规则的模式构建的序列预测的有效性。

## 2 实验及分析

本节详细介绍了图像序列数据集的构造过程及原则, 并在所构造数据集上利用深度神经网络对各维度的图像数字序列的预测加以分析比较。

### 2.1 数据集

正如本文前面所提及, 图像数字序列预测问题的解决方法与传统数字序列的解决方法不同, 它从另一种角度来解决数字序列预测问题。传统数字序列是在已知数字的意义下, 然后根据序列背景知识、结构复杂性等进行模式的构建, 从而求解数字序列, 而本文的方法是计算机在不知道如图 3 所示的  $a$ 、 $b$ 、 $c$ 、 $d$  以及  $A$ 、 $B$ 、 $C$ 、 $D$  间关系的前提下自动地学习出其中所包含的内在逻辑模式, 进而预测出所要求解的图像。

本文构造了 4 个大的数据集, 其包含各种图像序列, 在本文中, 根据其解决方案的不同将数据集分为 4 种不同的类型, 分别为: Linear、Multiplication、Fio 和 Nested。各类型图像序列的详细描述如下: 1) Linear 序列下一张图像的生成仅与前一图像内容相关联, 与其他项内容无关, 其仅涉及“+”和“-”两种数学运算, 例如等差序列就是 Linear 序列的一个特例; 2) Multiplication 序列仅包含“×”运算符, 并且其每张图像的生成原理与 Linear 序列的生成原理相同, 即仅与前一图像内容相关; 3) Fio 序列不仅与前一图像的内容相关联, 而且与前 2 张图像的内容相关联, 所涉及的运算符包括“+”和“-”。例如 Fibonacci 序列就是 Fio 序列的一个特例; 4) 对于 Nested 序列而言, 其包含更多的操作运算符: “+”、“-”和“×”, 并且其运算不仅仅与前 2 项相关, 还与任意的随机常数项相关联。用数学公式可以表述为:

$$f(n) = f(n-1) \pm f(n-2) \pm b \quad (3)$$

$$f(n) = [f(n-1) \pm f(n-2)] \times b \quad (4)$$

式中:  $f(n)$  表示当前图像;  $f(n-1)$  表示前一图像。式 (3) ~ (4) 均是 Nested 序列的数学表述。从另一层面来讲, 它可以说是 Multiplication 序列、Linear 序列和 Fio 序列的复杂混合序列。

本文整个数据集的构造原则: 1) 为方便构造数据集, 本文只考虑整数图像序列, 图像序列可以包含正数或负数; 2) 每张图像都包含一个多位数, 且每张图像中的最大位数设置为 10。所有这

些图像均是黑白图像,对齐和尺寸标准化(15×85);3)每张图像位数的设置最大为10位,本文选择的数据集仅是一个小样本集,其在总集(即我们数据的变动范围[-999 999 999 9, 999 999 999])中的占比很小,也就是说,本文确保训练集、测试集互不相交。本文为每种类型的数据集均选择了60 000的图像数字序列,其中50 000用于训练,10 000用于测试。Linear序列、Multiplication序列、Fio序列和Nested序列的部分数据集如图4~7所示(以4维的图像数字序列为例)。

538	1035	1532	2029
17	-326	-669	-1012
3812	3716	3620	3524
349	392	435	478

图4 Linear序列  
Fig. 4 Linear sequences

6257	43799	306593	2146151
-87	609	-4263	29841
-1	8	-64	512
94	1598	27166	461822

图5 Multiplication序列  
Fig. 5 Multiplication sequences

9	-132	-123	-255
489	253	742	995
-92	-43	-135	-178
4740	-205	4535	4330

图6 Fio序列  
Fig. 6 Fio sequences

506	77	-413	-1332
2660	726	-27088	210896
658	1264	2887	5116
-32	41	-180	2780

图7 Nested序列  
Fig. 7 Nested sequences

## 2.2 实验结果及分析

到目前为止,解决数字序列预测的方法几乎都是在基于给定数字的含义下,通过归纳序列的潜在规律,构建出满足该数字序列规律的模式,从而预测数字。而且,对于传统数字序列而言,其模式的长度与最终的预测结果有很大的关系,

即输入节点的数量对最终预测结果的准确率有极大的影响。Ragni等<sup>[18]</sup>提出了一种基于人工神经网络的动态学习方法来解决数字序列预测问题,该实验结果表明:输入节点的最佳配置为4个节点。因此,本文对输入维度,即输入的图像数进行了实验并加以分析比较。

本文对已构建的4个数据集进行了维度上的扩展,分别扩展为4维图像序列、5维图像序列和6维图像序列,这里并没有扩展3维图像序列,因为它的不确定性太大,例如本文的Fio序列,其并非简单地与最后一个数字有关,而是与最后2个数字相关联。关于数据集及其分类部分在3.1节中已有详细介绍,此处将不再赘述。

对于实验1,首先给定计算机4张图像,与此同时并不告知计算机每张图像中包含的内容是什么,在学习过程中,本文利用4种代表性的深度神经网络(MLP、LSTM、CNN\_MLP、ResNet)来完成数字序列逻辑学习任务,以便计算机可以从前3张图像间的逻辑模式自动学习出第4张图像。在测试阶段,本文任意输入4维图像序列,并将前3张图像放入模型中进行测试,使用OCR工具<sup>[27]</sup>来识别得到的预测图像,将OCR的结果与期望输出进行比较,并计算预测正确的百分比。

对于实验2,其步骤与实验1几乎相同。区别之处在于:在学习阶段,本文给计算机提供5张图像,让计算机自动学习每张图像间的内在逻辑模式。在测试阶段,输入5维图像序列,并将前4张图像传递到模型中以预测结果的正确性。

对于实验3,步骤与前2个实验几乎相同。不同的是,在学习阶段,本文给计算机提供6张图像,让计算机自动学习出每张图像间的逻辑模式。在测试阶段,本文输入6维图像序列,并将前5张图像传入模型进行测试,以检测本文是否可以生成具有正确结果的图像。

本文利用上述提及的4种深度神经网络模型对每个实验进行了准确度的测试,实验结果如表1所示,从表1可以看到4种神经网络模型可以很好地预测Linear序列和Fio序列,且维度越高,预测准确率越高。对Multiplication序列和Nested序列而言,MLP模型和LSTM模型对这2种序列的预测均呈现一种维度越高准确率越低的趋势。相反,CNN\_MLP模型和ResNet模型呈现一种维度越高,预测的准确率越高的情形,且ResNet模型

对于这2种序列的预测结果总体而言要优于其他的模型。

表1 4种神经网络模型在本文数据集上的实验结果  
Table 1 Experimental results on the our own dataset for the four neural network models

模型	输入维度	准确率/%			
		Linear	Fio	Multiplication	Nested
MLP	4	91.59	92.46	61.07	48.39
	5	92.08	93.48	50.81	47.74
	6	92.66	94.44	45.04	39.56
LSTM	4	91.75	92.23	71.51	50.79
	5	92.03	93.49	62.15	48.16
	6	93.35	95.1	50.3	42.38
CNN-MLP	4	89.26	90.57	54.38	48.28
	5	90.82	90.94	71.11	49.76
	6	91.21	94.55	85.13	51.89
ResNet18	4	89.54	90.43	64.09	51.3
	5	90.98	91.03	79.79	51.87
	6	91.93	93.4	86.5	53.21

本文对4种深度神经网络模型在4维、5维和6维图像序列的性能比较见图8~10。从各个维度图像序列的折线图中可以明显观察到,4种模型可以很好地预测Linear序列和Fio序列的结果,无论它们是4维、5维还是6维图像序列,且维度越高,准确率越高。这可能是由于Linear序列和Fio序列自身规律较为简单,例如,给定Linear图像序列5、7、9,预测结果11,计算机可以很明显知道是一个加2的规律,进而可以很容易地学到内在逻辑模式,从而进行序列的预测。

4种深度神经网络模型在4维图像序列的性能比较见图8, LSTM对Multiplication序列的预测准确率明显高于其它模型,但当维度升高时,如图9~10所示,可以看到LSTM模型对该序列的预测下降了,而ResNet和CNN\_MLP模型对于该序列的预测效果呈现越来越好的趋势,这可能是由于ResNet和CNN\_MLP网络模型较为复杂,当维度越高其获得的信息就会越多,越是可以较为准确的学到该图像序列的逻辑模式。而LSTM模型可能由于维度越高,其更易受到其子序列的影响。例如6维图像序列:1、2、4、8、16,预测32,该图像序列中包含子序列1、4、16和2、8等。在

LSTM模型中其极易受到子序列的影响,从而致使预测结果有误,如子序列1、4、16,其是一个乘4的规律,有可能下一个数预测为64,而当只给定4维图像序列时,如2、4、8,预测结果16,其反而较易进行该序列的预测。

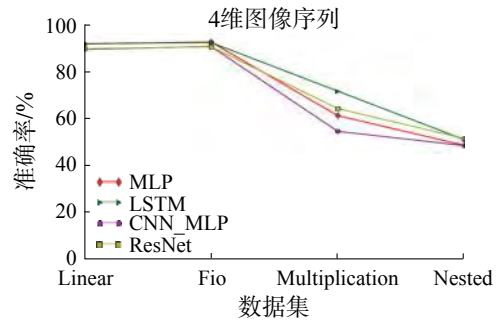


图8 4种深度神经网络模型在4维图像序列中的性能比较  
Fig. 8 A comparison on performance of four kinds of deep neural network models in 4-dimensional image sequences

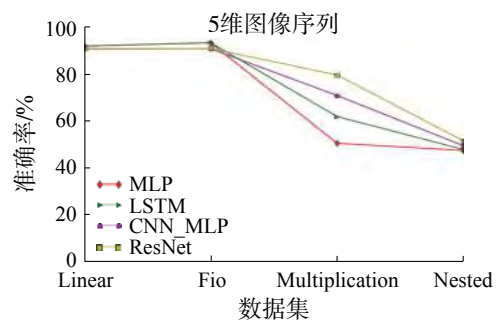


图9 4种深度神经网络模型在5维图像序列中的性能比较  
Fig. 9 A comparison on performance of four kinds of deep neural network models in 5-dimensional image sequences

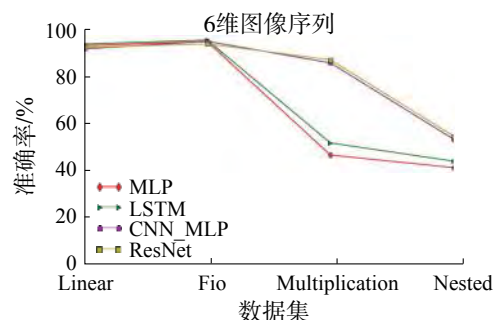


图10 4种深度神经网络模型在6维图像序列中的性能比较  
Fig. 10 A comparison on performance of four kinds of deep neural network models in 6-dimensional image sequences

对4维和5维图像序列的性能比较见图8~9,各模型对于Nested序列的预测均集中在一定的范围,这可能是由于Nested序列较为复杂,所有模型在该序列上的性能没有太大区别,但随着



维度增高,从图10可以发现 ResNet 和 CNN\_MLP 模型对于该序列的预测稍有提升,而 MLP 和 LSTM 模型对于此序列的预测又有所下降,这可能是由于 MLP 和 LSTM 模型较为简单,而 ResNet 和 CNN\_MLP 模型较为复杂,对于这种复杂序列,随着维数增高,复杂模型在图像序列的预测上其准确率极有可能上升,而简单模型的准确率极易下降。

此外,本文分析比较4种深度神经网络模型在 Linear 序列、Fio 序列、Multiplication 序列和 Nested 序列上的性能,如图11~14。从图11可以发现:随着维数的增高,各模型对于 Linear 序列预测的准确率提升并不明显,这可能是与 Linear 序列自身的规律有关,其规则是仅与前一项内容相关联,而与其他项无关,所以当维数增高时,计算机所获得的序列的内在逻辑模式的信息实则是没有太大的变化,例如,给定序列2、3、4,预测结果5,计算机可以很明显知道其是一个加1的规律,当再给定序列1、2、3、4预测下一个内容时,计算机并没有从中得到更多的信息,故而维度虽然增高了,但其准确率的变化并没有很大改变。

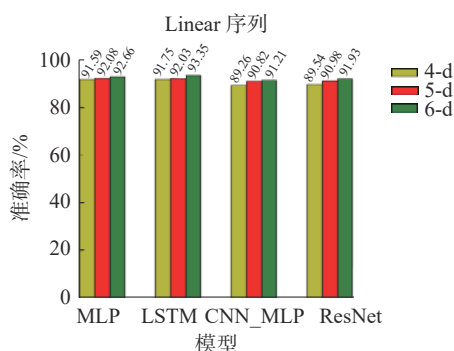


图11 4种深度神经网络模型在 Linear 序列上的性能比较  
Fig. 11 A comparison on performance of four kinds of deep neural network models in Linear sequences

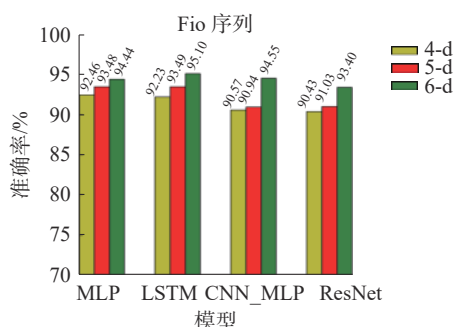


图12 4种深度神经网络模型在 Fio 序列上的性能比较  
Fig. 12 A comparison on performance of four kinds of deep neural network models in Fio sequences

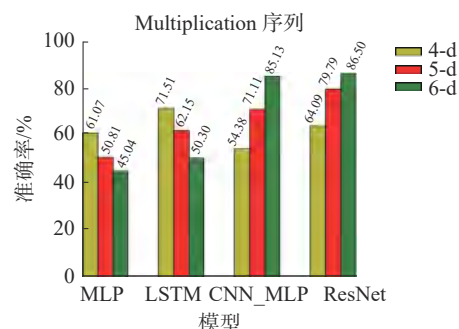


图13 4种深度神经网络模型在 Multiplication 序列上的性能比较

Fig. 13 A comparison on performance of four kinds of deep neural network models in Multiplication sequences

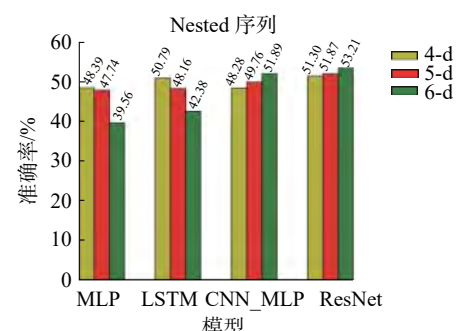


图14 4种深度神经网络模型在 Nested 序列上的性能比较

Fig. 14 A comparison on performance of four kinds of deep neural network models in Nested sequences

对比图12的Fio序列,其规则是不仅仅与前一项内容相关联,而且与前两项的内容相关,这就使得当维数越高,即给定的图像数越多,计算机得到的图像间的信息就会越多,就越易预测该序列,可以很明显地看到图12中准确率的提升随着维数的增高越来越明显。

Multiplication 序列和 Nested 序列的性能比较见图13~14,从图中可以看出维度越高,对于 CNN\_MLP 和 ResNet 模型而言预测的准确率越高。相反,MLP 和 LSTM 模型对于这2种序列的预测均呈现一种维度越高准确率越低趋势。这可能是由于 CNN\_MLP 和 ResNet 模型自身较为复杂,其对于这种复杂序列来说,其预测精度反而是易于随着维数增高而增大。而对于 MLP 和 LSTM 模型而言,其模型自身较为简单,加之 Multiplication 序列和 Nested 序列又较为复杂,相当于是个模型简单任务难的问题,其对于此序列的预测难度就会加大,容易导致维度越高准确率越低的情况。

## 2.3 结果分析

本小节从视觉效果的角度展示了每种类型图

像序列在不同维度上的实验结果,如图15~17所示(其中包括正确的预测和错误的预测)。

	Linear		Multiplication	
input	3899	562	-619	-5594
	3742	202	4952	22376
	3585	-158	-39616	-89504
output	3428	-500	316928	358016
ground truth	3428	-518	316928	358016
predict	correct	wrong	correct	wrong

	Fio		Nested	
input	488	776	1622	214
	-398	-468	1444	708
	90	308	3901	130
output	-308	-160	6180	86
ground truth	-308	-160	6180	46
predict	correct	wrong	correct	wrong

图15 4维图像序列

Fig. 15 4-dimensional image sequences

	Linear		Multiplication	
input	4261	444	-1575	-8125
	4620	224	-3150	24375
	4979	4	-6300	-73125
	5338	-216	-12600	219375
output	5697	-436	-25200	-658125
ground truth	5697	-436	-25200	-658125
predict	correct	wrong	correct	wrong

	Fio		Nested	
input	226	240	291	2175
	-144	-432	531	319
	82	-192	-88	2487
	-62	-624	-467	2799
output	20	-816	-1465	5209
ground truth	20	-816	-1465	5279
predict	correct	wrong	correct	wrong

图16 5维图像序列

Fig. 16 5-dimensional image sequences

本文仅展示了4维、5维、6维图像序列的部分实验结果。从这些实验结果发现预测误差的原因可能如下:1)当输出图像模糊时,很容易引起OCR工具识别错误,例如图15中的Fio序列,观察错误的那一栏,当输出图像模糊时,图像中输出结果的第2位(从右往左)有可能是6或8,这就极易导致OCR的识别出错,再如,图15~16中的Multiplication序列,图16中的Linear序列和Fio序列等。2)其对负号的图像预测有时是不准

确的,例如图17中的Multiplication序列和Nested序列等。3)结果输出中的位数越大,其预测出错的可能性就会越大,因为计算位数越大,相对而言,其难度就会越大,预测错误的可能性就会加大。

	Linear		Multiplication	
input	3528	4877	-2974	8682
	3715	4512	-17844	-60774
	3902	4147	-107064	425418
	4089	3782	-642384	-2977926
	4276	3417	-3854304	20845482
output	4463	3050	-23125824	-145318374
ground truth	4463	3052	-23125824	-145918374
predict	correct	wrong	correct	wrong

	Fio		Nested	
input	5267	557	2298	-212
	-634	-291	1886	-319
	4633	266	4379	-351
	3999	-25	6460	-490
	8632	241	11034	-661
output	12631	276	17689	1971
ground truth	12631	216	17689	-971
predict	correct	wrong	correct	wrong

图17 6维图像序列

Fig. 17 6-dimensional image sequences

### 3 结束语

数字序列逻辑学习任务一直以来都是人工智能的一个研究热点。而且未知的数字和规则的模型构建数字序列问题的预测始终是一个具有挑战性的任务。传统的数字序列技术是基于已知数字的含义下,并根据序列本身的规律特征构造满足该规律的模式,以便进行数字序列的预测。本文的方法是在不知道图像间关系和图像内包含的内容的意义的情况下利用深度学习的方法完成数字序列逻辑学习任务来预测下一张图像内容。同时,人类对于这种未知的数字和规则的模型构建的数字序列极易限制其中,无法快速计算出结果甚至无法得到正确结果。而本文所提出的方法为解决此类序列问题提供了一种可能。

在未来的工作中,我们将进一步探索如何提高精度及对缺失数据的预测,同时期望可以对任意位置的图像内容进行预测,而非单一地对下一张的图像内容进行预测,特别地,下一步希望可以提出更适用于数字预测逻辑学习任务的模型以求精度上面的突破。



## 参考文献:

- [1] BOGOSHI J. The oldest mathematical artefact[J]. *The mathematical gazette*, 1987, 71(458): 294–294.
- [2] SCHMID U, KITZELMANN E. Inductive rule learning on the knowledge level[J]. *Cognitive systems research*, 2011, 12(3/4): 237–248.
- [3] FLNER P, SCHMID U. An introduction to inductive programming[J]. *Artificial intelligence review*, 2008, 29(1): 45–62.
- [4] COLTON S, BUNDY A, WALSH T. Automatic Invention of Integer Sequences[M]// *The fertility of the soil* /. Cambridge University Press AAAI/IAAI, 2000: 558–563.
- [5] QIAN Y, LI F J, LIANG J Y, et al. Space structure and clustering of categorical data[J]. *IEEE transactions on neural networks and learning systems*, 2016, 27(10): 2047–2059.
- [6] QIAN Y, XU H, LIANG J Y, et al. Fusing monotonic decision trees[J]. *IEEE transactions on knowledge and data engineering*, 2015, 27(10): 2717–2728.
- [7] QIAN Y, LI Y, LIANG J, et al. Fuzzy granular structure distance[J]. *IEEE transactions on fuzzy systems*, 2015, 23(6): 2245–2259.
- [8] QIAN Y, LIANG J, PEDRYCZ W, et al. Positive approximation: An accelerator for attribute reduction in rough set theory[J]. *Artificial intelligence*, 2010, 174(9/10): 597–618.
- [9] RAVEN J, et al. Raven progressive matrices[M]. *Handbook of Nonverbal Assessment*, 2003, 223–237.
- [10] MUGGLETON S, CHEN J. Guest editorial: special issue on inductive logic programming (ILP 2011)[J]. *Machine learning*, 2012, 89(3): 213–214.
- [11] HOLZMAN T G, PELLIGRINO J W, GLASER R. Cognitive dimensions of numerical rule induction[J]. *Journal of educational psychology*, 1982, 74(3): 360–373.
- [12] STERNBERG R.J. *Handbook of Intelligence*[M]. Cambridge University Press, 2000.
- [13] SIMON H A, KOTOVSKY K. Human acquisition of concepts for sequential patterns[J]. *Psychological review*, 1963, 70(6): 534–546.
- [14] SANGHI P, DOWE D.L. A computer program capable of passing I.Q. tests[C]//4th Int. Conf. on Cognitive Science(ICCS'03), Sydney, 2003: 570–575.
- [15] SIEBERS M, SCHMID U. Semi-analytic natural number series induction[J]. *Advances in artificial intelligence*, 2012, 7526: 249–252.
- [16] STRANNEGARD C, AMIRGHASEMI M, ULFSBACKER S. An anthropomorphic method for number sequence problems[J]. *Cognitive systems research*, 2013, 22–23: 27–34.
- [17] HOFMANN J, KITZELMANN E, SCHMID U. Applying inductive program synthesis to induction of number series a case study with IGOR2[J]. *Joint german/austrian conference on artificial intelligence(Künstliche Intelligenz)*. Springer, Cham, 2014: 25–36.
- [18] RAGNI M, KLEIN A. Predicting numbers: An AI approach to solving number series[C]// *German Conference on Advances in Artificial Intelligence*. Springer, Berlin, Heidelberg, 2011:255–259.
- [19] RAGNI M, KLEIN A. Solving number series - architectural properties of successful artificial neural networks[C]// *NCTA 2011- Proceedings of the International Conference on Neural Computation Theory and Applications*, 2011: 224–229.
- [20] KINGMA D P, BA J. Adam: A method for stochastic optimization. [C]//*Proceedings of the 3rd International Conference on Learning Representations*, San Diego, arXiv preprint arXiv: 1412.6980, 2014.
- [21] LECUN Y, BENGIO Y, HINTON G. Deep learning[M]. *Nature*, 2015, 521(7553):436.
- [22] HOCHREITER S, SCHMIDHUBER, JÜRGEN. Long short-term memory[J]. *Neural computation*, 1997, 9(8): 1735–1780.
- [23] CHUMERIN N. Convolutional neural network[J]. *Neuro-computing*, 2015, 148: 136–142.
- [24] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. In: *advances in neural information processing systems*, 2012: 1097–1105.
- [25] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. *Journal of machine learning research*, 2014, 15(1): 1929–1958.
- [26] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[J]. In *proceedings of the IEEE conference on computer vision and pattern recognition*, 2016: 770–778.
- [27] SMITH R. An overview of the tesseract OCR engine[C]// *icdar*. IEEE Computer Society, 2007, 2: 629–633.

## 作者简介:



梁慧, 女, 1994 年生, 硕士研究生, 主要研究方向为机器学习、深度学习和逻辑学习。



曹峰,男,1980年生,副教授,博士,主要研究方向为人工智能、空间数据挖掘。主持国家自然科学基金项目1项,山西省青年科技研究基金项目1项,参与山西省青年科技研究基金项目2项,获中国科学院大学优秀毕业生称号,博士论文被评为中国科学院地理科学与资源研究所优秀博士论文。发表学术论文10余篇。



钱宇华,男,1976年生,教授,博士生导师,主要研究方向为人工智能、大数据、复杂网络、数据挖掘与机器学习。2014—2016年,连续入选爱思唯尔中国高被引学者榜单。曾获得山西省科学技术奖(自然科学类)一等奖,教育部宝钢教育基金特等奖,CCF优秀博士论文奖,山西省“五四青年奖章”,全国百篇优秀博士论文提名奖,获发明专利2项。发表学术论文80余篇。

## Thirty-seventh International Conference on Machine Learning (ICML 2020)

Submitted papers can be up to eight pages long, not including references, and up to twelve pages when references and acknowledgments are included. Any paper exceeding this length will automatically be rejected. Authors have the option of submitting one supplementary manuscript containing further details of their work and a separate file containing code that supports experimental findings; it is entirely up to the reviewers to decide whether they wish to consult this additional material.

To foster reproducibility, we highly encourage authors to submit code. Reproducibility of results and easy availability of code will be taken into account in the decision-making process. All submissions must be electronic, anonymized and must closely follow the formatting guidelines in the templates; otherwise they will automatically be rejected. This year, the author list at the submission deadline will be considered final, and no changes in authorship will be permitted for accepted papers.

会议日期: 2020-07-12 - 2020-07-17

截稿日期: 2019-12-06

会议地点: Messe Wien Exhibition & Congress Center, Vienna, 奥地利

网站: <https://icml.cc/Conferences/2020>