

# 利用稀疏编码结合深度学习的人体姿态估计

赵 康

(商丘职业技术学院, 河南 商丘 476100)

**摘 要:** 针对部分遮挡及模糊场景中的人体动作图像姿态估计精确不高的问题, 提出了一种基于稀疏编码和深度学习的人体姿态估计方法。首先, 对动作图像进行稀疏编码, 用较少的已知基函数重构测试样本。然后, 在训练算法中利用空间关系独立地分类像素, 进一步通过像素分类将图像分割为部分区域, 将前景区域分割加入到人体部位估计中。最后, 利用空间关系训练分类器实现人体部位估计的实时处理。在 CDC4CV 数据集上的验证结果表明, 提出的方法能够有效提取出高分类性能, 在许多情况中精度超过了 85%。

**关键词:** 稀疏编码; 部位估计; 像素分类; 深度学习; 人体姿态

**中图分类号:** TP391      **文献标识码:** A

## Human pose estimation using sparse coding and deep learning

ZHAO Kang

(Shangqiu Vocational and Technical College Shangqiu 476100 Henan Province China)

**Abstract:** Aiming at the problem that pose estimation of human motion images in partially occluded and fuzzy scenes is not accurate, a human pose estimation method based on sparse coding and deep learning is proposed. Firstly, the motion images are sparse coded, and the test samples are reconstructed with fewer known basis functions. Then, in the training algorithm, spatial relations are used to independently classify pixels, and the image is further segmented into part regions by pixel classification, and the foreground region segmentation is added into the estimation of human body parts. Finally, the spatial relation training classifier is used to realize the real-time processing of human body parts estimation. Verification results on CDC4CV dataset show that the proposed method can effectively extract high classification performance, and the accuracy exceeds 85% in many cases.

**Key words:** sparse coding; part estimation; pixel classification; depth learning; human pose

## 0 引言

人体姿态估计已经成为当前计算机视觉领域中一个重要的研究领域<sup>[1-2]</sup>, 具有非常重要的应用价值。例如, 人机交互、游戏、运动性能分析、身份鉴别等<sup>[3-4]</sup>。由此可见, 提出一种行之有效的姿态估计方法显得非常重要。

文献[5]提出将边缘特征和区域特征结合的新方法, 文献[6]提出预先利用人的检测先得到人体的大致区域, 再通过结合形状模型进一步进行更加细致的分割, 利用轮廓形状分类提高姿态估计准确率。文献[7]利用多特征融合的方法提高姿态估计

的准确率。然而, 当数据较多时, 这些方法容易出现过拟合现象, 且对于部分遮挡及模糊场景中的动作图像, 通常很难取得令人满意的结果。

基于此, 提出了利用稀疏编码结合深度学习的人体姿态估计方法。该方法的贡献之处总结如下:

①利用深度卷积网络非监督预训练的空间学习, 可提取出非常丰富的有效特征;

收稿日期: 2020-02-25

基金项目: 河南省高等学校重点科研项目(15A520118); 河南省科技厅软科学研究计划项目(142400411213)

作者简介: 赵康(1983-), 男, 硕士, 讲师, 研究方向为人工智能、图像处理等。

②利用空间关系将图像分割为部分区域,并将前景区域分割加入到人体部位估计中,从而有助于提高估计的准确率。

## 1 稀疏编码

在稀疏编码的数学模型理论中,基函数通过线性组合构成自然图像,并与自然图像的局部时频域特征相对应。当自然图像在某一方向或频率上存在最大特征时,与之对应的神经元会产生强烈的响应。所以当图像投影到由基函数所构成的特征子空间的时候,部分处于兴奋状态的神经元就产生相对于自然图像的稀疏表示。

稀疏编码数学模型的公式可表示为:

$$I(x, y) = \sum_i a_i \Phi_i(x, y) \quad (1)$$

其中,  $I(x, y)$  为自然图像的灰度函数;  $\Phi_i(x, y)$  为第  $i$  个基函数;  $a_i$  为第  $i$  个神经元的响应, 可以看成基函数对当前图像的贡献值。另外, 稀疏编码还考虑了一个局部特征描述子和少量特征单词之间的关联。该模型可以表示为以下优化问题:

$$\arg \min_{A, x_i} \{ \|y_i - Ax_i\|^2 + \lambda \|x_i\|_1 \} \quad (2)$$

式中,  $y_i$  表示局部特征描述子;  $A$  表示特征字典;  $x_i$  表示  $y_i$  对应的稀疏系数。通过特征提取构造特征字典  $A = [a_1, a_2, \dots, a_M]$ 。对于待识别目标  $y$ , 可以利用字典  $A$  求解出对应的稀疏系数。图 1 为面向空间目标识别的稀疏编码过程。

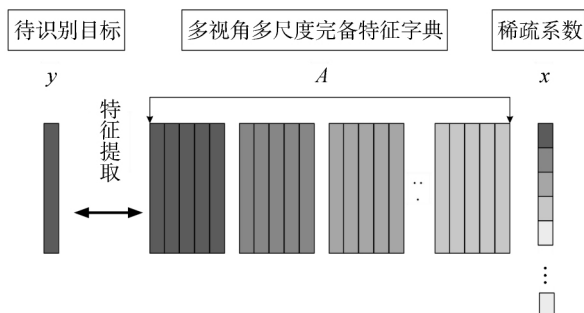


图1 面向空间目标识别的稀疏编码过程

## 2 融合空间约束的深度学习

在上一节,通过对动作图像进行稀疏编码,用较少的已知基函数重构测试样本。在这一节,提出了一种将图像分割为子块的方法,本文将这个任务佯装成典型的分类问题,在该分类中,将图像的每个像素分类为部分  $L = \{1 \dots L\}$  之一。本文旨在决策像素,以最小化计算复杂性。

本文给定  $M$  个图像  $\{X^1, \dots, X^M\}$  及其相关的已标记的地面实况。在本文方法中,图像的像素被类似地索引为指数:  $X^M = \{X^M(i)\}$ 。通过不断学习由以

下两部分组成的分割模型。

用于将每个图像  $X$  嵌入到一个特征表示的参数映射函数  $Z = f(X | \theta_f)$ 。该函数由  $Q$  个映射组成,每个映射的维度与  $X$  的相同,因此可以被线性索引:  $Z(i) \in R^Q$ 。从训练数据中学习参数  $\theta_f$ , 这期间在嵌入空间中考虑了特征对间的欧几里得距离:  $d(i, j) = \|Z(i) - Z(j)\|_2$ 。

分类器  $\hat{y}(i) = g(Z(i) | \theta_g)$ , 该分类器从给定训练参数  $\theta_g$  中分类特征  $Z(i)$ , 并给出了部分区域标签的估计  $\hat{y}(i)$ 。

新方法假设部分区域的邻接信息因图像不同, 接下来的两小节将分别描述如何将空间受限整合到嵌入  $Z = f(X | \theta_f)$  的训练过程和分类器  $\hat{y}(i) = g(Z(i) | \theta_g)$  的训练过程。可以独立应用或组合这两个过程。对于每个元素  $i$ , 将相邻布局上的信息注入到独立运行的分类器上。

### 2.1 空间预训练

学习嵌入函数  $Z = f(X | \theta_f)$  的常用方法是最小化非监督自动编码环境中<sup>[8]</sup>或监督环境中的重建误差, 并将具有相同标签的像素映射到特征空间中附近的点, 将不同标签的像素映射到远点。

考虑到真实数据的可用性(即部分标签), 本文在预训练阶段侧重使用。文献[9]已经呈现了监督或半监督训练, 以生成更有识别力的嵌入。然而, 本文并不是将分类损失用作训练标准。本文引入了一个新的能量函数, 该能量函数利用不同地区的空间布局:

$$E = \sum_{i, j} \delta_{y(i), y(j)} L_s(i, j) + \sum_{i, j} v_{y(i), y(j)} L_D(i, j) \quad (3)$$

式中,  $y(i)$  和  $y(j)$  分别表示由  $i$  和  $j$  索引的像素的地面实况标签;  $v_{a, b}$  的定义是: 如果  $a$  和  $b$  不是语料库中的邻居, 则  $v_{a, b} = 1$ , 否则为 0。在特征空间中, 在由  $i$  和  $j$  索引的特征对之间测量  $L_s$  和  $L_D$ , 该过程是由原始数据  $X$  通过内嵌函数嵌入的。本文忽略来自具有不同标签对的贡献, 但是它们各自部分在语料库中是相邻的。因为邻域之间没有明确界限, 所以本文忽略单个边界上的特征之间的关系。

基于保留空间关系设计能量函数还有许多其他方法。例如通过量化标签之间的距离(例如身体部分之间的平均距离), 则可将这种“软”邻居标准并入能量函数。

$L_s$  为具有相同标签的像素对的损失组件, 它起到推动其嵌入的作用。 $L_D$  为具有不同标签的像素对的对比损失组件, 它起到将其嵌入分开的作用。类似

于文献[11]  $L_s$  和  $L_D$  的设置:

$$L_s(i, j) = \frac{1}{2} (d(i, j))^2 \quad (4)$$

$$L_D(i, j) = \frac{1}{2} [\max(0, \alpha - d(i, j))]^2 \quad (5)$$

其中,  $\alpha > 0$  为界限, 它控制着对比范围。

通过链式法则(反向传播)可以计算方程(1)中对应参数的梯度:

$$\frac{\partial E}{\partial \theta_f} = \frac{\partial E}{\partial Z(i)} \frac{\partial Z(i)}{\partial \theta_f} + \frac{\partial E}{\partial Z(j)} \frac{\partial Z(j)}{\partial \theta_f} \quad (6)$$

由方程(4)和(5)可以获得  $\frac{\partial E}{\partial Z(i)}$  和  $\frac{\partial E}{\partial Z(j)}$ 。

$\frac{\partial Z(i)}{\partial \theta_f}$  和  $\frac{\partial Z(j)}{\partial \theta_f}$  取决于嵌入函数  $Z = f(X | \theta_f)$  的特定体系结构。

尽管能量函数是在成对像素方面上定义的, 但是用像素进行测试, 所以测试不需要解决组合问题。

## 2.2 监督空间 LR(Logistic Regression) 学习

本文选择具有单全连接分类层(即上层相当于逻辑回归)的卷积网络, 该卷积网络是端到端的训练。相同原则在像素分类中引入了空间关系, 也可以将其应用到其他分类器上。特别地, 在前期研究中, 基于最大化熵增加, 本文提出了一种通过适应训练算法而在学习随机森林中包含空间受限的方法。

训练基于分类的神经网络主要是为了最小化交叉熵。当网络的归一化输出被当作概率时, 这相当于最大化了网络分配给真正类的日志概率。在多元环境中, 可以将它们与地面实况标签相比较来归一化网络的输出。然而, 最小化交叉熵时不考虑部分标签的布局。

本文提出了以下新的损失函数, 它们根据网络输出基于类标签的排名。对于每个输入向量, 正向传递给出每个标签的网络响应, 它可用以按降序排序类标签, 可根据这种排序和期望的目标排序之间的差异定义损失, 用以下属性定义此损失:

①排名最高的类标签应为目标地面实况标签, 这种约束与传统神经网络学习中的熵损失有关。

②排名第二的标签应为语料库类邻域定义中的地面实况标签。本文侧重通过迫使网络学习这些约束, 实现更好的泛化到看不见的数据。

对于以上属性的实例如图2所示, 图中像素的地面实况标签为1, 实际输出地面实况标签排为第2。目标标签排序将地面实况标签1排为第1位, 接着是标签3和2, 在这个实例中, 标签3和2为标签1的邻居。

学习排序是机器学习中的一个传统问题, 本文

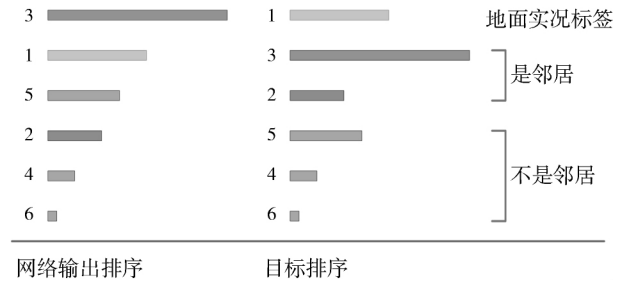


图2 基于排序差异的本文提议的损失函数

采用在成对约束上的定义。给定标签对  $(u, v)$ , 用  $g_u(Z(i))$  和  $g_v(Z(i))$  分别表示位置  $i$  的最上(即分类)层输出, 用  $o_{uv}(i) = g_u(Z(i)) - g_v(Z(i))$  表示它们的偏差。通过逻辑函数计算排序比  $v$  高的标签  $u$  的概率:

$$P_{uv}(i) = \frac{e^{o_{uv}(i)}}{1 + e^{o_{uv}(i)}} \quad (7)$$

给定以上描述的期望排序的属性, 推导出了以下两组成对约束:

①  $L-1$  个成对约束集, 这里, 每个约束规定, 地面实况标签的排序高于任一其他标签。

②对于一组约束, 每个约束规定, 属于地面实况标签邻居的标签  $u$  排序高于其他标签  $v$ , 其中, 标签  $v$  不是地面实况标签的邻居。

单一像素的损失函数  $\varepsilon$  为成对损失  $C_{uv}$  的成对约束的总和:

$$\varepsilon(i) = \sum_{u \neq 1} C_{1u}(i) + \sum_{(u, v): v \neq 1} C_{uv}(i) \quad (8)$$

其中,  $i$  表示像素的指数, 本文将简式  $l = y(i)$  用作其地面实况标签。基于排序的损失函数提供了结合分类损失和空间布局有原则的方式。通过控制目标

概率  $\hat{P}_{uv}$  的  $\lambda$  参数可以设置两类约束。改变这个参数, 可以将不同的属性给不同的约束。

然而该方法允许将未标记数据包括到半监督环境中的学习过程中。在这种情况下, 将利用以上描述的损失函数将标记图像程序给网络, 然而, 未标记数据的损失函数不包含包括了地面实况标签的成对约束, 本文考虑了网络最强输出的邻近, 而没有考虑地面实况邻近。

## 3 提出方法的详细过程

尽管映射  $Z = f(X | \theta_f)$  可能是任意可调架构, 但本文采用两阶段卷积网络, 每个阶段都包含具有一堆过滤器的卷积层、空间池和二次抽样操作符, 在与过滤器卷积之后, 生成一堆特征映射。第一阶段的输出为第二阶段卷积层的输入, 这里将每个特征映射连接到之前阶段的几个映射。在学习开始之前

均匀随机地选择连通性。对于网络输出,计算能量函数的梯度,同时采用反向传播更新参数 $\theta_f$ 。参数 $\theta_f$ 由特征映射处的过滤器、特征图上的乘法偏差组成。

由于不同部分复杂交互,许多视觉识别问题需要大的上下文信息。然而,卷积网络的上下文大小受限于过滤器尺寸和采样率。简单地选择更大的过滤器并未解决在不同尺度上可能观察到重要线索的问题。为了克服这个难题,在本文体系结构中采用多尺度卷积网络。本文使用 $f(\theta_s^f)$ ,  $\forall s \in \{1, \dots, N\}$ 表示由每个尺度卷积网络产生的输出,这里 $s$ 表示尺度, $N$ 为尺度的总数。采用不同尺度公有的权值,卷积网络可以学习尺度不变特征。

在预处理阶段,为每个像素构造多尺度拉普拉斯金字塔 $X^s$ ,  $\forall s \in \{1, \dots, N\}$ ,其中 $X^1$ 具有与初始映射一样的尺寸。将局部对比度归一化应用于拉普拉斯图像,以确保局部邻域的平均值和标准偏差为零。从每个尺度图像 $X^s$ 上随机提取由补丁对组成的批处理。利用相应的卷积网处理这些补丁,这里所说的参数 $\theta_f^s$ 为不同的尺度公用。

以上结构的训练过程包含两个步骤,第一步骤为空间深度学习,该步骤中的标签只用于定义空间布局。本文因此称这个阶段为有监督的预训练。在这个阶段,初始化卷积网络参数,以致这些特征与标签的空间排列一致。第二步为有监督的空间学习。为了预估标签,将 $\theta_g$ 参数化的逻辑回归层连接到最上层卷积网络的特征映射。本文还应用了一种改进方案,在该方案中,反向传播来自LT的梯度以更新卷积网络参数 $\theta_f$ 。

提取出来的两个高斯过滤响应之间的差异体现在形状索引的高斯差分(Difference of Gaussian, DoG)特征上,最后计算预定义的DoG过滤的响应。将相似性转换应用到局部视网膜采样模式,形状索引的DoG特征计算并不包含整个图像的转换,只转换稀疏的坐标。另外,获取高斯平滑图像的计算复杂性并不高,因为平滑过程只有一次,且在学习过程之前。

用单一的面部坐标来学习高斯过程回归树(Gaussian Process Regression Tree, GPRT)树<sup>[10]</sup>,由随机采样阈值的DoG特征获取树的分割函数。为了获取更有识别力的分割函数,测试了一些分割函数并选取了性能最佳者。而分割函数的性能由第1面部坐标的平方损失来衡量。这一过程仅学习了树的分割函数,而回归输出由GPRT学习。

最初阶段的学习树倾向于使用形状索引的DoG特征,这些特征由距离较远的采样点计算得来,而处

于后期的树倾向于使用从附近的采样点计算得来的特征。这是因为较远的采样点设计更大的区域,对于形状评估误差来说更稳定。附近的采样点稳定性差一点,但是当形状评估准确时识别力高。因此,在学习阶段,形状索引的DoG特征允许每一个树根据实际需要自由选择更稳定的特征。

## 4 实验结果

关于来自深度图像的特定人体部分估计的应用,为了在不同情形中用多尺度卷积网络验证空间学习,本文对提出的框架进行了实验验证。

本文在以下所有实验中使用二阶段卷积网络,第一个卷积层由8个尺寸为 $9 \times 9$ 的过滤器组成,接着是元素级正切函数和 $2 \times 2$ 非重叠平均池算子。第二个卷积层由32个 $9 \times 9$ 的滤波器组成,每个滤波器用相同的激活和池算子整合前一层4个特征映射,其输出为32个特征映射。对于单一尺度上的单个卷积网,其感受域为 $28 \times 28$ 。在具有3个尺度金字塔且每个尺度包含2个卷积层的多级环境中,感受域尺寸为 $112 \times 112$ ,所以使用大的空间背景学习更好的特征。

将局部对比度归一化用于卷积网络的输入中,同时,将来自背景的像素设置为任意高的值(在本文情况中为4)以区分目标的零分布像素。预训练的边限 $\alpha$ 设为 $1.25^{[11]}$ ,对于所有实验的所有约束,将 $\lambda$ 设为定值1,即如果 $u$ 排序高于 $v$ ,则 $\bar{P}_{uv} = 1$ ,否则为0。弱监督特征学习和监督学习每个通过数据集使用30个时刻,这里采用小批量梯度下降。如果采用端到端微调,则使用另外30个时刻。依靠经验选取不同的学习超参数(例如不同层的学习率),例如,对于第一个卷积层,将 $\varepsilon_1$ 设为 $10^{-6}$ ,对于第二个卷积层,将 $\varepsilon_2$ 设为 $10^{-5}$ 。当进行分类时,本文考虑每个孤立的像素,将它所相应的32-维特征向量用于LR的输入。在每个像素位置公用LR的参数。在下面的实验中,本文将介绍平均像素精度。

文献[12]设计了CDC4CVPoselets数据集,以解决上半部人体分割从转化的、重新尺度化的深度图像到灰度域(0-255)的问题。该数据集包含基于kinect骨架模型的上身姿势的345个训练和347个测试深度图像,背景差分已经预处理了所有图像。这些身体部分为:头部、颈部、上臂、前臂、左臀部、臀部。此外,给出腰部以下部分其他类的标签。在从标准姿势计算的11个部分之间,本文定义了28个邻近关系的布局,在标准姿势中,主体延伸他或她的手臂,例如,头-颈、颈-左肩膀、颈-右肩膀、左前

臂-左上臂,等等。在空间预训练的三个不同尺度随机选择 900 对补丁。

对于来自深度图像的部分估计的应用,本文的基线是文献[13]中描述的随机决策森林算法,和在没有任何空间背景的情况下,具有监督的端到端训练的单尺度和多尺度卷积网络。特别的,实验验证了文献[5]提出的边缘特征与区域特征相结合的方法、文献[7]采用的多特征融合的方法。实验还将本文框架与随机决策森林中的最近空间学习算法<sup>[14]</sup>进行了比较。在本文的基线实验中,研究了深度为 9 的 3 棵树的随机森林。对于每个节点,从 4000 个偏置候选者和 22 个阈值候选者中选择最佳偏置和阈值。更多细节请见文献[14]。图 3 是 CDC4CV 数据集上<sup>[15]</sup>几种不同方法的对比情况。结果表明,提出的方法均优于几种对比方法。此外,在 CDC4CV 数据集上,基于深度学习的体系结构学性能优于随机森林。

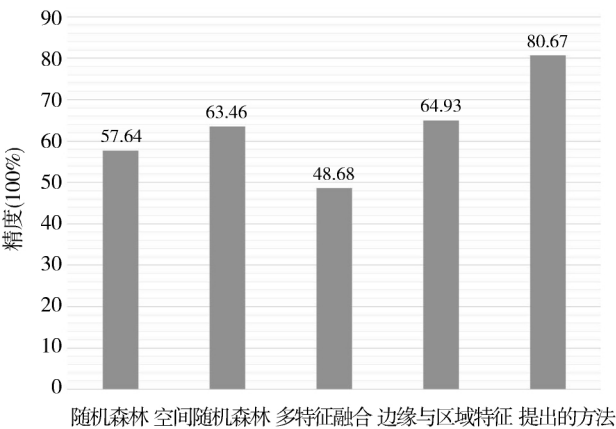


图3 CDC4CV 数据集上不同方法的比较

基于纯卷积层(称作卷积网络)和分类层(称作 LR)的空间预训练,研究了不同框架结合的性能。并依照非空间基线,实施了 DrLIM 的像素版,一个类似预训练的策略,在该策略中并不考虑空间布局。结果如表 1 所示。对于每个设置(微调或不微调),空间训练的性能优于非空间训练,且在许多情况中,精度很高。这是因为提出的方法在框架网络的监督学习中整合了空间部分区域分布。能以极小的成本获得空间学习的增益,并提高了分类器的性能。

表1 CDC4CV 数据集上的结果

卷积层	LR	精度( % )
DrLIM	传统的	66.67
DrLIM	空间深度学习的	85.76
空间深度学习的	传统的	73.43
空间深度学习的	空间深度学习的	86.45

续表 1

卷积层	LR	精度( % )
DrLIM	传统的	65.46
DrLIM	空间深度学习的	86.32
空间深度学习的	传统的	66.41
空间深度学习的	空间深度学习的	87.95

5 结束语

本文提出了基于稀疏编码和深度学习新方法来提高分割问题中的分类性能。提出的算法中的能量函数不包括成对像素关系,降低了计算成本。相比于测试阶段中同一类型的标准像素分类器,本文方法的复杂度增加值为零。实验结果证明了提出方法的精确性,通过初始阶段将图像、视频或它们的组成图像分割为部分区域可以解决许多计算机视觉领域的问题。本文方法集成了部分区域的空间布局,且不需要添加额外的步骤到像素分类中。且性能优于基于随机决策森林的著名方法和其他有监督的特征映射方法。

参 考 文 献:

[1] 毕雪超. 基于空间骨架时序图的舞蹈特定动作识别方法[J]. 信息技术 2019 43(11):16-19 23.

[2] 曾志超 李桂清 邹歆仪 等. 三维人体模型姿态与形状重构[J]. 计算机辅助设计与图形学学报 2019 31(9):1485-1493.

[3] Szczuko P. Deep neural networks for human pose estimation from a very low resolution depth image[J]. Multimedia Tools and Applications 2019 78(5):29357-29377.

[4] 刘希. 基于深度卷积神经网络的图像语义分割[J]. 信息技术, 2019 43(2):71-74.

[5] Dehghankhalili M ,Fallahi S ,Mahmudi F ,et al. Epidemiology ,regional characteristics ,knowledge ,and attitude toward female genital mutilation/cutting in southern iran[J]. Journal of Sexual Medicine , 2015 12(7):1577-1583.

[6] 彭莱. 基于主动形状模型和 POEM 纹理模型的人脸特征点定位方法研究[D]. 重庆: 重庆邮电大学 2016.

[7] You Y ,Zhao Y. A human pose estimation algorithm based on the integration of improved convolutional neural networks and multi-level graph structure constrained model [J]. Personal and Ubiquitous Computing 2019 23(7):607-616.

[8] 杨兴明 周亚辉 张顺然 等. 跨阶段结构下的人体姿态估计[J]. 中国图象图形学报 2019 24(10):1692-1702.

[9] 李世宝 王升志 刘建航 等. 基于接收信号强度非齐性分布特征的非监督学习室内定位指纹库构建[J]. 电子与信息学报, 2019 41(10):2302-2309.

[10] 李键红 吴亚榕 吕巨建. 基于自相似性与多任务高斯过程回归的单帧图像超分辨率重建[J]. 光学精密工程 2018 26(11):2814-2826.

[11] Lin Chien-ming ,Tsai ,Chi-yi ,et al. Visual object recognition and pose estimation based on a deep semantic segmentation network [J]. IEEE Sensors Journal 2018 18(22):9370-9381.

(下转第 69 页)

供更低的 MAE 和 RMSE 分数。

为了获取所实现的基于常规模型的 CF 方法的更多信息,表 4 给出了所提出方法的混淆矩阵。

表 4 所提出方法的混淆矩阵

		实际				
预测	评分	1	2	3	4	5
	1	214	54	27	16	5
	2	279	668	333	161	40
	3	571	1245	4256	3430	1213
	4	91	256	1016	3584	2584
	5	0	0	0	1	1

精确度、召回率、 $F$  值和准确度以及产生结果所需的迭代次数如表 5 所示。

表 5 基于模型的 CF 的精度比较

模型	精确度	召回率	$F$ 值	准确度	迭代次数
NMF	0.49	0.33	0.39	0.44	12
MLP	0.46	0.32	0.37	0.42	17
MF	0.44	0.28	0.34	0.31	25

一般来说,就评分预测而言,基于模型的 CF 往往比基于记忆的 CF 表现更好。表 5 表明,与其他基于模型的 CF 相比所提出的 NMF 模型在精确度、召回率、 $F$  值和准确度方面性能更优。应该注意的是,本文提出的模型在迭代次数上也少于另外两个模型,因此计算效率也更好。所提出的模型在速度和性能方面比其他常用的协作过滤方法更好。

## 4 结束语

本文探讨了基于深度神经网络和矩阵分解的反馈推荐模型。矩阵分解模型与神经网络体系相结合,明显提高了推荐模型的性能。在实验中,提议的模型为未评分项目提供了接近实际评分的预测,这可能会对电子商务、娱乐和个性化推荐产生现实影响,从而提供更好的用户体验和销售增长。该模型

仅使用了用于预测的用户、项目和评分,未使用用户统计信息(即年龄、职业)等变量,因此在未来的研究中将更改用户和项目嵌入的影响因素的数量,进一步提升模型的性能和鲁棒性,并对模型的性能进行更加深入的评估。

## 参考文献:

- [1] 杨凯,王利,周志平,等.基于内容和协同过滤的科技文献个性化推荐[J].信息技术,2019,43(12):11-14.
- [2] 李振波,杨晋琪,岳峻.基于协同回归模型的矩阵分解推荐[J].图学学报,2019,40(6):983-990.
- [3] 王运,倪静,马刚.基于 FunkSVD 矩阵分解和相似度矩阵的推荐算法[J].计算机应用与软件,2019,36(12):245-250.
- [4] 艾静超.基于改进协同过滤技术的个性化旅游线路推荐研究[J].现代电子技术,2019,42(23):182-186.
- [5] 张家精,夏翼鹏,陈金兰,等.基于张量分解和深度学习的混合推荐算法[J].南京大学学报:自然科学,2019,55(6):952-959.
- [6] 胡思才,孙界平,据生根,等.基于深度神经网络和概率矩阵分解的混合推荐算法[J].四川大学学报:自然科学版,2019,56(6):1033-1041.
- [7] 黎雪微,应时,周寅.基于用户兴趣迁移的网络图书推荐模型研究[J].图书馆学研究,2019(22):56-65.
- [8] 厉小军,柳虹,施寒潇,等.基于深度学习的课程推荐模型[J].浙江大学学报:工学版,2019,53(11):2139-2145,2162.
- [9] 曹银,李伟.基于信任和图聚类的协同过滤推荐算法[J].信息技术,2019,43(10):77-82.
- [10] 邱神芬,黄业文,宋杰,等.基于深度矩阵分解网络的矩阵填充方法[J].计算机科学,2019,46(10):55-62.
- [11] 邓存彬,虞慧群,范贵生.融合动态协同过滤和深度学习的推荐算法[J].计算机科学,2019,46(8):28-34.
- [12] 宋瑞雪,李国勇.基于改进的矩阵分解模型在推荐系统中的应用[J].计算机应用,2019,39(S1):93-95.
- [13] 王永贵,尚庚.融合注意力机制的深度协同过滤推荐算法[J].计算机工程与应用,2019,55(13):8-14.
- [14] 李同欢,唐雁,刘冰.基于深度学习的多交互混合推荐模型[J].计算机工程与应用,2019,55(1):135-141.
- [15] 吴彦文,刘闯.基于用户偏好和可疑度的推荐方法研究[J].计算机应用研究,2018,35(12):3632-3634.

责任编辑:杨静

\*\*\*\*\*  
(上接第 65 页)

- [12] Holt B, Ong E J, Cooper H, et al. Putting the pieces together: connected poselets for human pose estimation [C]. IEEE International Conference on Computer Vision Workshops. IEEE, 2011: 1196-1201.
- [13] Hilde E P, Sicco A B, Frans N, et al. Gait patterns in association with underlying impairments in polio survivors with calf muscle weakness [J]. Gait & Posture, 2017, 58: 146.

- [14] Andrew J J, Smith, Edward D L. Temporal-spatial gait parameter models of very slow walking [J]. Gait & Posture, 2018, 61: 125-129.
- [15] Hansen L, Siebert M, Diesel J, et al. Fusing information from multiple 2D depth cameras for 3D human pose estimation in the operating room [J]. International Journal of Computer Assisted Radiology and Surgery, 2019, 14(8): 1871-1879.

责任编辑:丁玥