

参赛队员姓名: 罗雨诗

中学: 清华大学附属中学

省份: 北京

国家/地区: 中国

指导教师姓名: 贾珈

论文题目: 基于深度学习的服装美学建模

本参赛团队声明所提交的论文是在指导老师指导下进行的研究工作和取得的研究成果。尽本团队所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果。若有不实之处，本人愿意承担一切相关责任。

参赛队员：罗雨诗

指导老师：贾琳

2017年9月15日

论文题目：基于深度学习的服装美学建模

论文摘要：服装是当下社会潮流的标志，围绕服装细节识别和分类的研究依旧是计算机领域的研究热点。现有的研究课题进一步将服装细节特征与服装的视觉效果及美学风格进行关联，实现了自动的服装美学的鉴赏。本文将基于已有的研究的细节特征数据和美学词汇，利用深度学习的思想来构建二者之间的关联模型，力求提升模型的有效性和准确性。我们计算出美学评价词和种子词汇之间的距离，将这些美学词汇定位到二维美学坐标中，建立了美学评价空间。之后我们引入深度学习的思想构建了服装细节特征与美学风格之间的关联建模，并设计了对比实验证明了该模型的有效性。在研究过程中，我们探讨了深度学习中的不同参数和结构对关联建模结果的影响，通过对比实验选择了最佳的深度学习模型参数。在上述工作的基础上，我们还分析了电商数据中不同类别服装风格的差异，进一步验证了模型的有效性。

关键词：服装美学鉴赏，美学语义空间，深度学习

Title: Aesthetic modeling based on Deep Learning

Abstract: Clothing is one of the symbol of modern fashion trends, research concerning clothing detail's recognition still remains a hot topic in the field of computer. Existing research is mainly focused on relating clothing detail features and aesthetic styles, realizing the possibilities of auto clothing aesthetic appreciation. In this paper, based on existing data sets, we will introduce the ideology of Deep Learning into the association model between clothing visual features to improve the model's effectiveness and accuracy. By calculating the distance between aesthetic words and seed words, we are able to map the most often used aesthetic words in a two-dimensional space, building a fashion semantic space. Later, we introduce the ideology of Deep Learning into the association model between clothing detail features, and designed experiments to prove the effectiveness of the model. In the research, we discussed how different parameters and structures in Deep Learning effect the outcome of the relating model, and by comparison we choose the best parameters of the model. Based on the work mentioned above, we analyze different styles of different categories' clothing in shopping websites, which further confirms the effectiveness of our model.

Keywords: clothing aesthetic appreciation, fashion semantic space, deep learning

# 目录

<b>1</b>	<b>引言</b>	<b>1</b>
<b>2</b>	<b>相关工作</b>	<b>2</b>
2.1	服装图片识别及分类的相关研究 . . . . .	2
2.2	基于服装特征的美学风格研究 . . . . .	3
<b>3</b>	<b>问题定义</b>	<b>4</b>
3.1	美学语义空间的定义 . . . . .	4
3.2	研究方法 . . . . .	4
<b>4</b>	<b>数据及特征</b>	<b>4</b>
4.1	数据库的总体描述与示例 . . . . .	4
4.2	对特征标注的详细解释 . . . . .	5
<b>5</b>	<b>基于深度学习的服装美学鉴赏建模</b>	<b>6</b>
5.1	DNN简介 . . . . .	7
5.2	DNN在美学鉴赏中的建模 . . . . .	7
5.2.1	形式化表述 . . . . .	8
5.2.2	模型结构与训练 . . . . .	8
<b>6</b>	<b>实验与讨论</b>	<b>9</b>
6.1	实验数据 . . . . .	9
6.2	实验设置 . . . . .	9
6.2.1	DNN中不同参数和结构对关联建模结果的影响 . . . . .	9
6.2.2	对比DNN和SVM模型的优劣 . . . . .	10
6.3	实验指标 . . . . .	10
6.4	实验结果与讨论 . . . . .	10
6.4.1	深度学习中参数实验结果的影响 . . . . .	10
6.4.2	对比DNN和SVM模型的运行结果 . . . . .	11
6.5	案例分析 . . . . .	11
<b>7</b>	<b>总结</b>	<b>13</b>
<b>A</b>	<b>附录</b>	<b>15</b>

# 1 引言

随着日常生活中科技与经济的快速发展, 人们对物质生活的要求随之提升, 大多数人开始更加注意自己的形象。穿着打扮作为其中非常重要的一部分, 受到了很多人的重视。服饰风格作为一个人内在的外化显示, 可以体现一个人的修养和性格, 选择合适的服装便是日常生活中必不可少的一部分。然而, 服装种类十分多样且繁杂, 想要能够从大量的服装中选出适合自己的衣服, 仅仅靠观察和手动筛选是很低效的。因此, 利用能够迅速处理大量数据的计算机来实现服装审美, 便是一个较好的解决办法。让计算机学会去理解服装的美, 去自动地欣赏并评价美是几年来逐渐被人重视的研究课题, 也是本文希望能研究的目标。然而, 要让计算机理解何为“美”并不是一件易事。一般来说, 计算机是通过程序和算法模拟人脑的逻辑思维, 而人类对美学的认识其实更处于感性认识层面, 如何计算出这类感性认知, 是非常困难的。研究如何让计算机学会欣赏美将会是非常有意义的工作, 而如果赋予计算机以欣赏美的能力, 大众追随时尚潮流会变得更加便捷, 人们的生活品质将会得到进一步提高, 同时也让我们更进一步地对人工智能的可能性进行深入探索。

针对服装图片的分析, 国内外已经有许多关于服装分类识别的研究。Wei Yang[1]构建了一个可以识别服装细节特征的应用系统, Kota Yamaguchi[2]研究如何对比筛选出风格类似的服装, Masaru Mizuochi[3]构建了一个由局部检索整体的服装检索系统。但是上述研究工作都仅限于服装细节的识别, 而并未将这些特征与服装的视觉效果及美学风格进行关联。最新的数篇研究从让计算机感性地理解美入手, 拓展了计算机自动分析和识别服装的美学效果的研究领域。Jia Jia[4]构建了服装视觉特征-图像度量空间-美学词汇空间三层框架, 建立起视觉特征与美学词汇之间的联系。Yihui Ma[5]构建了一个规模庞大的服装图片特征数据集, 并设计了可以对服装搭配进行美学评价的模型。这些研究工作在服装细节特征和美学效果关联建模的模型主要应用了自编码器模型 (autoencoder<sup>1</sup>)。本文我们将借鉴Yihui Ma[5]发布的公开数据集, 利用深度学习的思想来对服装细节特征和其美学效果进行关联建模, 力求实现对服装美学效果更准确的自动鉴赏。

本文将主要研究如何在已有的服装细节特征数据和美学词汇之间, 引入深度学习的思想, 来实现二者之间的关联建模。在数据层面上, 我们采用Yihui Ma[5]文章中所构建的数据库, 其中包含了43596张从专业的流行时尚网站vogue.com上截取的近十年来时装秀图片还有60004张从电商网站jd.com上截取的包含不同样式的共服装商品图片; 基于该数据库和其所提供的服装图像的细节特征标注, 我们利用Deep Neural Network (DNN)模型来对服装细节特征和美学效

---

<sup>1</sup>[http://deeplearning.stanford.edu/wiki/index.php/Autoencoders\\_and\\_Sparsity](http://deeplearning.stanford.edu/wiki/index.php/Autoencoders_and_Sparsity)

果进行关联建模，实现了一个从特征到美学之间的映射。为了证明该模型的有效性，我们设计了多组参数对DNN的关联建模结果影响的实验，分别对迭代次数、隐含层数和神经元数进行了实验，并进一步通过对比实验筛选出最佳的参数组合，并与经典的回归模型SVM运行的结果进行对比。实验表明采用DNN的方法在迭代次数为300、共六层隐含层、每一层神经元数均为80的实验指标之下的运行结果，可以将误差值从0.2083降至0.1862，证明了该模型的有效性和准确性。

本文的创新点总结如下：

- 将深度学习的思想引入到服装特征和美学效果的关联建模中，实现更为准确的服装美学自动评价；
- 探讨了深度学习模型中的不同参数和结构对关联建模结果的影响，通过对比试验选择了最佳的深度学习模型取值，较SVM的结果将准确率提升了10%左右；
- 基于上述模型在典型的电商数据上对服装美学进行量化描述，揭示了不同类别服装之间的风格差别。

## 2 相关工作

现代计算机研究中已经存在许多针对服装识别分类相关的研究，因此下面我们通过服装图片识别及分类的相关研究，以及基于服装特征的美学效果的相关研究这两方面，对国内国外的相关工作进行调研和介绍。

### 2.1 服装图片识别及分类的相关研究

Wei Yang[1]通过构建一个较为全面的应用系统模型，来解析服装图片并生成精准的服装标注。模型首先将图片进行切分，创造出不同的区域，然后把这些区域与不同元素进行了关联，从而构建出一个服装解析模型。研究表明，这种系统的可以较为准确地切割图片并识别服装。Kota Yamaguchi[2]从一个标注好的服装图片数据集中对比筛选出与指定服装图片风格类似的，通过大量的样例学习并识别服装细节元素。实验证明，这种检索方法较大程度的改善了在服装细节标注和解析方面的准确性。Masaru Mizuochi[3]构建了一个由局部特征检索服装的服装检索系统，通过关联大量图片中提取的局部特征于全局特征之间的关联，这种系统实现了根据服装整体和局部关联性检索类似服装款式。该研究将用户检索服装的满意程度提高了将近20%。Kevin Lin[6]构建了分层的深度搜索框架，第一层框架中是服装的视觉表达学习得到的预训练网络模型，第二层增加了隐含层来学习表

达哈希式, 第三层模块利用哈希编码来达到检索速度的提升。该检索框架在对大规模服装图片数据集的搜索中表现良好。李福娣[7]采用深度学习中的深度卷积神经网络结构进行特征学习, 将传统方法中图像美学评价特征作为数据输入的图层, 以此提升系统分类性能。实验表明, 该方法中的两个美学评价数据库中都取得良好的分类性能。

上述研究均取得了较好效果, 都能够从服装图片中较为准确的提取出服装细节特征。这些研究为我们在基于服装特征的基础上进行更多服装识别研究提供了技术支持。但是上述研究工作都仅仅限于服装细节的识别, 而并未将这些特征与服装的视觉效果及美学风格进行关联。

## 2.2 基于服装特征的美学风格研究

为了将服装视觉效果与美学风格进行关联, 一些研究围绕这种联系进行了深入探究。Jia Jia[4]在服装视觉特征和美学评价词汇之间引入中间层来建立联系的思想, 定义了视觉特征-图像度量空间-美学词汇空间三层框架, 建立起视觉特征与美学词汇之间的联系。实验表明, 这种三层式的框架可以更好地建立特征与美学词汇之间的联系。Yihui Ma[5]构建了一个规模庞大的服装图片特征数据集, 并设计了可以自动进行服装搭配并评价分析的模型, 融入了上下装之间搭配的隐含规则, 更好地实现了对服装搭配的美学鉴赏。Yejun Liu[8]设计了一款互动型的美学评价应用系统“魔镜”, 将美学词汇与视觉特征之间构建关联使得计算机学会自动欣赏美学风格。该系统可以基于时尚潮流针对服装的美学层面进行评价和服装推荐, 研究表明系统对于服装风格的鉴赏准确性良好。Jingtian Fu[9]在Yejun Liu[8]设计的“魔镜”基础上更进一步的改良改善该应用系统的准确性并满足更多用户的需求, 在时尚潮流的基础上增加了对于个人偏好的分析, 能够更加有针对性的对用户进行服装搭配推荐, 通过遗传算法该系统实现了对于用户的品味和喜好的学习。

上述研究拓展了计算机自动分析和识别服装的美学效果的研究领域。本文中我们将借鉴上述研究工作中所使用的数据及特征的公开资源。但是上述研究工作中所使用的将服装细节特征和美学效果关联建模的模型主要应用了自编码器模型, 并没有研究DNN模型在该场景的关联建模中的应用效果。所以本文中我们将在Yihui Ma[5]研究中的数据和特征的基础上, 将DNN应用到服装细节特征和其美学效果的建模中, 力求实现更准确的利用计算机对服装美学效果的自动鉴赏。

## 3 问题定义

如果我们要实现计算机对服装美学效果的计算，我们该如何去量化描述服装美学？为了去更好的来撰写相关数据和我们所提出的模型，我们首先对本文服装特征和美学效果的关联建模给出一个明确的定义。

### 3.1 美学语义空间的定义

在本文中，我们将参考 Jia Jia[4]的研究工作采用日本色彩心理学家 Shigenobu Kobayashi[10]的二维美学坐标。二维美学坐标是一个基于服装美学空间的二维空间冷暖(warm-cool)、软硬(hard-soft)。针对日常生活中服装的美学评价词，例如青春的、随性的等等，我们利用 Shigenobu Kobayashi[10]所给出的种子词汇及其在二维平面中的分布，通过WordNet::Similarity[11]来计算服装的美学评价词和种子词汇之间的语义距离，从而将这些美学词汇定位到二维美学坐标中。

### 3.2 研究方法

我们准备从三方面开展研究工作：

- 探究国际上最新的针对服装美学的公开数据库，来自Yihui Ma[5]的研究；
- 建立美学评价空间，借鉴 Jia Jia[4]的研究工作中筛选出的美学词汇相似的定义在二维美学空间坐标上；
- 在上述工作的基础上，将DNN模型引入到服装细节特征和美学词汇之间的关联建模。

## 4 数据及特征

### 4.1 数据库的总体描述与示例

为了将服装细节特征和美学评价词之间进行关联建模，我们需要一个标注特征种类全面且数据规模庞大的服装图片数据库。因此本文章采用Yihui Ma[5]研究中发布的一个国际公开的服装美学标准数据库来进行探究。

该数据库利用从专业的流行时尚网站vogue.com上截取的近十年来43596张时装秀图片，其中包括了大约1:1的男女服装，且大部分图片都同时包含上下装，示例见图1；另外还有从电商网jd.com上截取的包含不同样式的共60004张服装商品





图 1: 时装秀图片示例

图片, 其中男女装比例约为1:1, 且大部分图片都只包含单件上装或下装。图片数  
 据集中对服装特征的标注所制定的标注方案, 参见表1。

类别	特征	维数	类别	特征	维数	
上装	性别	3	裤子	性别	3	
	衣长	5		长度	4	
	袖长	5		腰部	3	
	领口	8		纹理	9	
	开口	7		材质	9	
	版型	5		版型	3	
	纹理	9		属性	9	
	材质	9		裙子	长度	3
	属性	15			版型	2
	其他	场合	8	纹理	9	
气候		3	材质	9		
调性		2	褶皱	2		

表 1: 特征标注方案

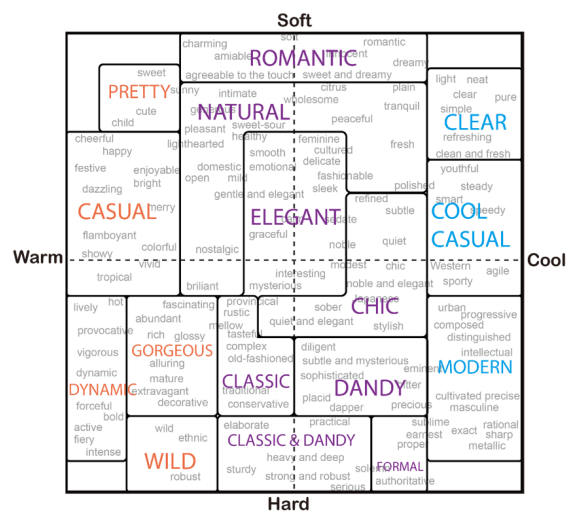


图 2: 美学二维评价空间

## 4.2 对特征标注的详细解释

具体地, 上装的性别特征被分为三个维度分别为男、女、通用; 衣长分为短款、中款、长款、露脐、超长; 袖长分为吊带、无袖、短袖、中袖、长袖; 领口分为毛领、立领、V领、一字领/大开领、圆领、翻领、高领、有连衣帽; 开口分为单排纽扣、半排纽扣、双排纽扣、套头、拉链、半排、拉链、敞开; 版型分为紧身、直筒、宽松、斗篷型、收腰; 纹理分为纯色、格子、圆点、碎花、横条纹、竖条纹、数字+字母、图案集中、图案重复; 材质分为混纺、针织、丝绸/雪

纺/纱、皮革、牛仔、纯棉、化纤、亚麻、呢绒；属性分为西装、运动服、冲锋衣、羽绒服、衬衫、毛衣、T恤、卫衣、大衣、皮衣、夹克、马甲、皮草、连衣裙、旗袍。类似地，裤子，裙子以及其他的每一标注项均如此，详细的服装特征标注请见附录A。需要特别说明的是，属性为服装的类别特征，如毛衣，衬衫等，不属于服装的基本特征；调性，即服装的美学效果，分别对应到二维平面上的一个点，详见图2。马依慧[12]研究中招聘兼职在标注网站上进行远程标注，系统给出对服装图片特征的问题，让标注者选择认为符合的选项。为了保证标注结果的准确性，每张图片每个特征由不同的人标注三次，最终标注结果由三个标注结果投票决定。

图片示例			
上装颜色			
上装样式	【性别】女 【衣长】(未知) 【袖长】长袖 【领口】翻领 【开口】单排纽扣 【版型】直筒 【纹理】纯色	【性别】女 【衣长】短款 【袖长】中袖 【领口】V领 【开口】单排纽扣 半排 纽扣 套头 【版型】直筒	【性别】男 【衣长】短款 【袖长】中袖 【领口】圆领 【开口】套头 【版型】直筒 【纹理】纯色
下装颜色			
下装样式	【长度】短裙 【版型】A字裙 【纹理】纯色 【材质】混纺 丝 绸/雪纺/纱 皮革 纯棉 化纤 亚麻 呢绒 【褶皱】无	【长度】七八分 【腰部】(未知) 【纹理】纯色 【材质】混纺 牛仔 纯棉 亚麻 呢绒 【版型】紧身 【性别】女	【长度】长裤 【腰部】(未知) 【纹理】碎花 图案重复 【材质】混纺 纯棉 【版型】直筒 【性别】男

图 3: 图片标注示例

根据上述的标注特征和标注方法，马依慧[12]将vogue.com上近十年来43596张时装秀图片和jd.com上截取的包含不同样式的共60004张服装商品进行标注，标注示例如图3。基于上述数据库，本文将进行基于深度学习的服装美学鉴赏建模。

## 5 基于深度学习的服装美学鉴赏建模

上面介绍了数据库中细节特征的标注，本章介绍深度学习模型的原理和在本研究中的应用。

## 5.1 DNN简介

DNN(Deep Neural Network) , 又称深度神经网络, 是在输入和输出之间学习一个映射关系的模型, 包含多层神经元, 每一层先进行权重的线性计算再经过激活函数, 最终可根据特定的输入向量得到对应的输出向量。DNN内部的神经网络层可以分为三类, 输入层、隐含层和输出层, 如图4所示<sup>2</sup>, 一般来说第一层是输入层, 最后一层是输出层, 而其余的中间层即为隐含层。

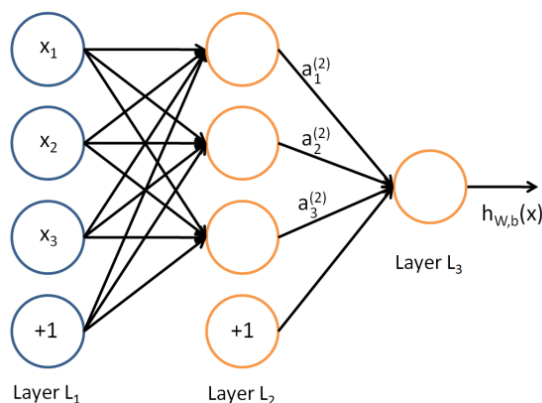


图 4: DNN基本结构

## 5.2 DNN在美学鉴赏中的建模

本节中将用形式化的语言介绍DNN在美学鉴赏任务中的建模方法, 表2中列出了出现的所有数学符号说明。

符号	说明	符号	说明
$V$	服装图片数据集	$h_i^{(l)}$	$v_i$ 对应的第 $l$ 层的向量
$v_i$	$V$ 中的第 $i$ 张图片	$W^{(l)}$	第 $l$ 层和第 $l+1$ 层之间的权重参数
$x_i$	$v_i$ 的服装视觉特征向量	$b^{(l)}$	第 $l$ 层和第 $l+1$ 层之间的偏移参数
$N$	$x_i$ 向量的维数	$\hat{w}c_i$	DNN输出的对于 $wc$ 的估计值
$Y$	美学二维平面	$J$	模型训练中的损失函数
$wc$	warm-cool, $Y$ 中的第一维	$m$	训练时输入的图片样本数
$hs$	hard-soft, $Y$ 中的第二维	$\lambda_1, \lambda_2$	$J$ 中的超参数
$M$	( $V \rightarrow X$ )到 $Y$ 的预测模型	$\theta$	梯度下降算法中的参数表示
$N_h$	DNN的隐含层数	$\alpha$	梯度下降算法中的步长大小

表 2: 符号说明

<sup>2</sup>[http://ufdl.stanford.edu/wiki/index.php/Neural\\_Networks](http://ufdl.stanford.edu/wiki/index.php/Neural_Networks)

### 5.2.1 形式化表述

给定服装图片数据集 $V$ , 对于其中每张图片 $v_i \in V$ , 我们可以用一个 $N$ 维向量 $x_i = \langle x_{i1}, x_{i2}, \dots, x_{iN} \rangle (\forall x_{ij} \in [0, 1])$ 来表示 $v_i$ 的上下装视觉特征, 用 $X$ 来表示一个 $|V| * N$ 维的特征矩阵,  $X$ 中的元素 $x_{ij}$ 表示 $v_i$ 的第 $j$ 维视觉特征。

此外, 我们用 $Y(wc, hs) (\forall wc, hs \in [-1, +1])$ 来表示美学二维平面, 其中 $wc$ 从-1到+1的变化对应着平面中从warmest到coolest的变化,  $hs$ 从-1到+1的变化对应着平面中从hardest到softest的变化。在 $Y$ 中分布了一系列带有坐标的人们常用来描述服装风格的词 (如文雅的、运动的、自然的等)。

我们所研究的问题是对于一张输入的服装图片, 找到它在 $Y$ 中对应的二维坐标点, 并根据点的位置对应到与其距离最近的风格描述词, 作为输入图片的服装风格。因此, 我们需要训练一个预测模型 $M : (V, X) \Rightarrow Y$ , 来完成从服装图片及视觉特征到美学二维平面的映射。

### 5.2.2 模型结构与训练

对于 $Y(wc, hs)$ 的两个维度, 我们对每一维训练一个单独的DNN模型, 并将其分别输出的一维结果组合起来, 作为二维平面中的坐标, 来完成从服装视觉特征到美学二维平面的映射。由于两个模型结构的相似性, 我们以warm-cool维度为例, 详细介绍DNN的层间关系以及模型的训练过程。

假设DNN一共包含 $N_h$ 个隐含层, 我们定义相邻层之间的关系如下:

$$h_i^{(l+1)} = \tanh(W^{(l)}h_i^{(l)} + b^{(l)}) \quad (1)$$

其中 $h_i^{(l)}$ 表示 $v_i$ 对应的第 $l$ 层的向量,  $W^{(l)}$ 和 $b^{(l)}$ 分别是第 $l$ 层和第 $l+1$ 层之间的权重和偏移参数,  $\tanh$ 为双曲函数 ( $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ )。特别地,  $h_i^{(0)} = x_i$ ,  $\hat{w}c_i = h_i^{(N_h+1)}$ 。

对于输出层的结果 $\hat{w}c_i$ , 我们用以下损失函数来衡量它与真实值之间的误差:

$$J(W, b) = \frac{\lambda_1}{2m} \sum_{i=1}^m \|wc_i - \hat{w}c_i\|^2 + \frac{\lambda_2}{2} \sum_l (\|W^{(l)}\|_F^2 + \|b^{(l)}\|_2^2) \quad (2)$$

其中 $m$ 是输入的图片样本数,  $\lambda_1, \lambda_2$ 是超参数, 用来控制两项之间的相对重要性,  $\|\cdot\|_F$ 表示Frobenius范数。第一项表示 $\hat{w}c$ 与 $wc$ 之间的误差, 第二项表示对权重和偏移参数的正则化, 用来减小它们的值以防止过拟合。

对于训练过程, 我们用 $\theta = (W, b)$ 来表示训练希望得到的参数。DNN的优化目标是寻找一组特定参数 $\theta^*$ , 使得损失函数 $J(W, b)$ 的值最小:

$$\theta^* = \arg \min_{\theta} J(W, b) \quad (3)$$

我们使用的优化方法是随机梯度下降算法[13]。对于每一轮迭代，参数的更新如下：

$$W = W - \alpha \frac{\delta}{\delta W} J(W, b) \quad (4)$$

$$b = b - \alpha \frac{\delta}{\delta b} J(W, b) \quad (5)$$

其中 $\alpha$ 是梯度下降算法中的步长大小。

将两个维度分别的对应DNN按照上述步骤训练完成后，我们的模型就可以利用最终得到的参数，对于输入的任意服装视觉特征 $x$ ，能够计算模型的输出 $wc$ 和 $hs$ ，将它们组合成 $y(wc, hs)$ ，即可对应到美学二维平面中的一个点，进一步可以对应到相关的美学词汇。

## 6 实验与讨论

### 6.1 实验数据

本节中，我们将对比深度学习模型和普通机器学习回归算法的实验结果，来验证深度学习模型的准确性和有效性。我们将采用第4章中介绍的数据库进行实验。

### 6.2 实验设置

针对本文进行实验中所涉及到的特征学习，我们采用DNN作为深度学习的模型，SVM为基础的机器学习模型。SVM是一种向量支持机，在过去几十年里SVM对分类和回归问题都有较好的实验效果，是一种非常经典的算法，因此我们选择SVM作为基础回归算法的代表。

实验分为两个步骤进行，首先分别通过改变迭代次数、隐含层数和神经元数来探究参数变化对DNN结果的影响，然后比较DNN结果和SVM结果，从而证明模型的有效性。接下来我们将具体介绍实验的设置。

#### 6.2.1 DNN中不同参数和结构对关联建模结果的影响

- 迭代次数对实验结果影响：首先固定神经元层数，改变迭代次数，每次进行五次实验取平均值，比较MSE值。迭代次数以50次为单位增长，共得到六组数据。
- 层数对实验结果影响：取实验结果最佳的迭代次数，固定该次数，增加神经元层数，每层进行五次实验取平均值，比较MSE值。每次增加一层，共得到

六组数据。

- 神经元数对实验结果影响：固定层数，增加每层神经元数。每组进行三次实验取平均值，比较MSE值。每次增加一层，共得到20组数据。

### 6.2.2 对比DNN和SVM模型的优劣

根据上述实验得出DNN的最佳实验结果后，我们将DNN和SVM的实验结果来表明深度学习模型和普通机器学习模型的实验结果对比。本文中的DNN和SVM模型分别使用了Python的开源代码（TensorFlow<sup>3</sup>和scikit-learn<sup>4</sup>）。

## 6.3 实验指标

本文实验中的所有实验值，我们均采用均方误差（Mean Squared Error, MSE）来计算模型的误差值。MSE值越小说明模型的误差越小，从而证明模型的准确性更高。最终的实验结果均为五折交叉验证的结果。MSE的计算公式如下：

$$MSE = \frac{1}{n} \sum_{i=1}^n (observed_i - predicted_i)^2 \quad (6)$$

## 6.4 实验结果与讨论

### 6.4.1 深度学习中参数实验结果的影响

在这节中我们将报告上述实验设置下DNN中的不同参数和结构对关联建模结果的影响。迭代次数对实验结果影响见图5，我们分别尝试了训练步数为50，100，150以此类推的迭代次数。由图可见，从50次到100次的迭代次数变化对MSE有较大影响，大幅下降；从100次到150次变化较小，有了些微的下降；而从200到300次之间的MSE并无明显区别，趋于平稳。从50次到300次实验值有了较为明显的变化，误差降低了3.80%，由此可见在迭代次数为300左右的时候可取的较为理想的实验数据，因此我们选取300次作为后续实验的得带次数标准值。

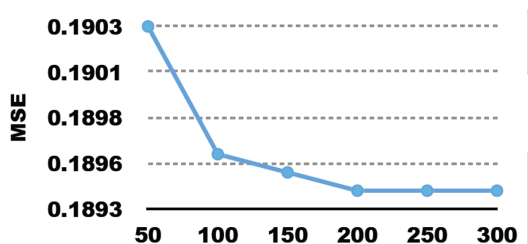


图 5: 迭代次数对关联建模结果的影响

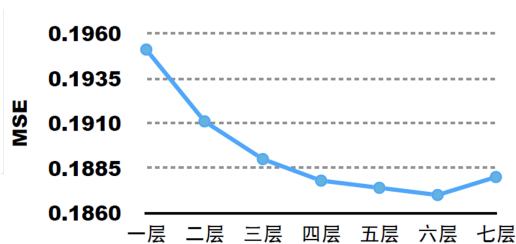


图 6: 层数对关联建模结果的影响

<sup>3</sup><https://www.tensorflow.org/>

<sup>4</sup><http://scikit-learn.org/stable/>

实验结果影响见图6，我们分别尝试了一层到七层。由图可见，从一层到二层MSE值约降低了2.05%，从二层到三层降低了1.09%，从三层到四层降低了0.06%，从四层到六层变化也相对较小，但是在六层到七层的时候实验数据反而向上增长，这说明层数过多可能会导致实验误差增大，因此我们选择六层作为后续实验的最大层数。

神经元数对实验结果影响见表3，在层数增长的基础上增加神经元数，我们发现当层数一定时随着神经元数的增加，MSE值逐渐减少，变化值在4%-8%之间，而当神经元数固定时，随着层数增加，MSE值呈现先下降，后升高的趋势。对比不同组的实验结果，我们选取迭代次数为300，共六层隐含层，每一层神经元数均为80的运行情况下得到的0.1862的运行结果。

	10	20	40	80
两层	0.1954	0.1915	0.1897	0.1872
三层	0.1952	0.1923	0.1881	0.1877
四层	0.1930	0.1899	0.1886	0.1875
五层	0.1928	0.1900	0.1883	0.1875
六层	0.2040	0.1917	0.1893	0.1862

表 3: 层数和每层神经元个数对关联建模结果的联合影响

#### 6.4.2 对比DNN和SVM模型的运行结果

实验可得DNN在本文探究的参数影响下可取的最佳实验值为0.1862，最差实验值为0.2040，而SVM的运行结果为0.2083，误差值均高于DNN的运行结果，最佳值甚至高出将近10.6%，远远好于SVM的运行结果。由此可见，通过参数调整后的DNN模型可以更好地对服装细节特征和美学效果之间建立联系，深度学习模型的有效性远远高于基础的机器学习模型，因此本文所提出的引入深度学习思想的关联建模可以实现更准确的关联建模。

### 6.5 案例分析

利用上述建立的模型，我们进一步针对电商数据中的服装风格进行了一些有趣的案例分析。具体地，我们从电商网站上爬取了著名运动品牌Nike<sup>5</sup>的100张服装商品图片，包括不同性别/年龄、不同运动种类、不同季节、上装或者下装等各种类型，然后将每一张图片映射到了美学二维平面中的一个点。进一步地，我们对服装的不同类型在二维平面中进行了风格对比。

<sup>5</sup><https://nike.tmall.com/>

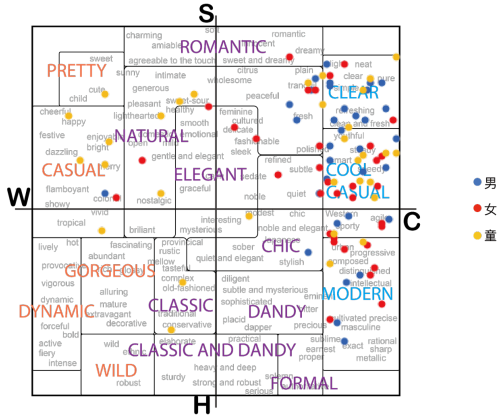


图 7: 不同性别/年龄服装对比

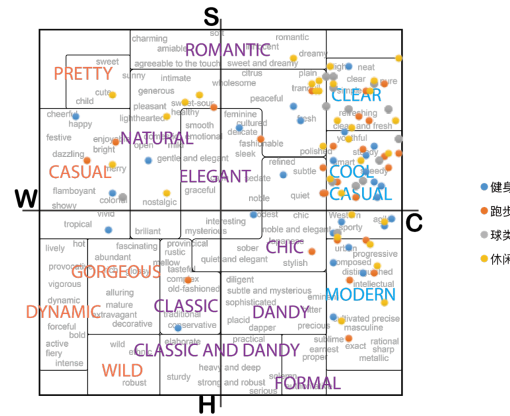


图 8: 不同运动种类服装对比

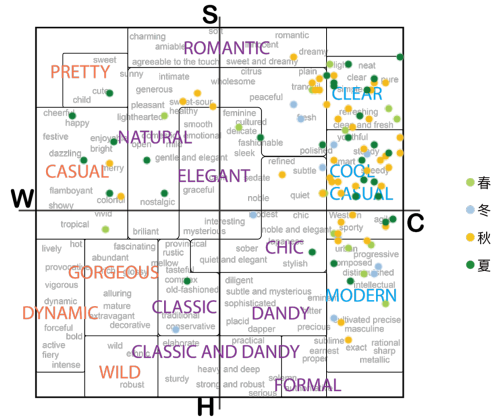


图 9: 不同季节服装对比

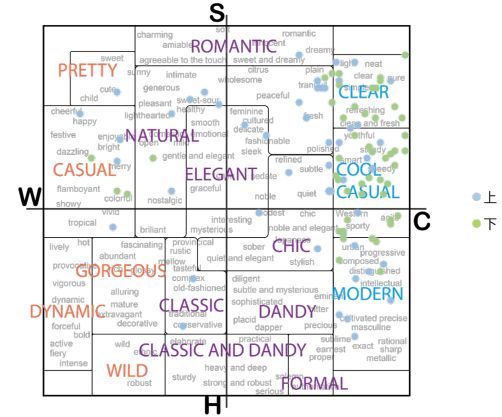


图 10: 上装和下装对比

- 不同性别/年龄服装对比：从图7中可以看出，男装分布地比较集中，基本上都处在平面最右侧的位置，对应着clear、cool、casual等。而女装的风格稍微向左有所拓展，除了casual、modern之外，还有些分布在elegant、natural等靠中间的区域。至于童装，则大致覆盖了整个平面，说明其风格十分多样化。
- 不同运动种类服装对比：从图8中可以看出，球类的服装分布地更为集中，而健身、跑步、休闲这三类服装较为分散，不仅仅局限于最右边一块区域，有的甚至分布在最左边，这说明这些类别中的服装存在一些较为独特的设计款式。
- 不同季节服装对比：从图9中可以看出，对于不同季节的服装，它们的风格也存在着一定的差异。比如，春秋装的分布十分相似，冬装的分布较为集中，而夏装由于其设计上的多样化，在分布上也是覆盖了平面上的多个区域，具有多种多样的风格。
- 上装和下装对比：从图10中可以看出，上装分布的范围远大于下装，这也与



我们的日常认知相符，即上装往往能设计出更多样更丰富的款式和色彩，而下装（尤其是类似于Nike这种运动品牌的下装）则风格比较单一。

## 7 总结

服装是人们一直关心的话题，围绕服装美学的研究是十分有意义的。本文中，我们基于已有的研究的细节特征数据和美学词汇，利用深度学习的思想来构建二者之间的关联模型，力求提升模型的有效性和准确性。首先，我们通过计算美学评价词和种子词汇之间的距离，将美学词汇定位到二维美学坐标中，建立了美学评价空间。然后，我们引入深度学习的思想构建了服装细节特征与美学风格之间的关联建模，并设计了对比实验证明了该模型的有效性。最后，我们还分析了电商数据中不同类别服装风格的差异，用案例分析进一步验证了该模型在此问题中的有效性。

## 参考文献

- [1] Wei Yang, Ping Luo, and Liang Lin. Clothing co-parsing by joint image segmentation and labeling. In *Computer Vision and Pattern Recognition*, pages 3182 – 3189, 2015.
- [2] K Yamaguchi, M. H. Kiapour, L. E. Ortiz, and T. L. Berg. Retrieving similar styles to parse clothing. *Pattern Analysis & Machine Intelligence IEEE Transactions on*, 37(5):1028–40, 2015.
- [3] Masaru Mizuochi, Asako Kanezaki, and Tatsuya Harada. Clothing retrieval based on local similarity with multiple images. In *ACM International Conference on Multimedia*, pages 1165–1168, 2014.
- [4] Jia Jia, Jie Huang, Guangyao Shen, Tao He, Zhiyuan Liu, Huanbo Luan, and Chao Yan. Learning to appreciate the aesthetic effects of clothing. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.
- [5] Yihui Ma, Jia Jia, Suping Zhou, Jingtian Fu, Yejun Liu, and Zijian Tong. Towards better understanding the clothing fashion styles: A multimodal deep learning approach. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- [6] Kevin Lin, Huei Fang Yang, Jen Hao Hsiao, Jen Hao Hsiao, and Chu Song Chen. Rapid clothing retrieval via deep learning of binary codes and hierarchical search. In *ACM on International Conference on Multimedia Retrieval*, pages 499–502, 2015.
- [7] 李福娣. 基于深度卷积神经网络的图像美学评价. PhD thesis, 华南理工大学, 2016.
- [8] Yejun Liu, Jia Jia, Jingtian Fu, Yihui Ma, Zijian Tong, and Zijian Tong. Magic mirror: A virtual fashion consultant. In *ACM on Multimedia Conference*, pages 680–683, 2016.
- [9] Jingtian Fu, Yejun Liu, Jia Jia, Yihui Ma, Fanhang Meng, and Huan Huang. A virtual personal fashion consultant: Learning from the personal preference of fashion. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

- [10] Shigenobu Kobayashi. The aim and method of the color image scale. *Color Research & Application*, 6(2):93–107, 1981.
- [11] Ted Pedersen, Siddharth Patwardhan, and Jason Michelizzi. Wordnet: similarity - measuring the relatedness of concepts. In *National Conference on Artificial Intelligence*, pages 1024–1025, 2004.
- [12] 马依慧. 基于美学原理的服装搭配研究与系统构建. Bachelor thesis, 清华大学, 2016.
- [13] Léon Bottou. *Large-Scale Machine Learning with Stochastic Gradient Descent*. Physica-Verlag HD, 2010.

## A 附录

<b>上装</b>	<b>性别</b>	男	女	通用						
	衣长	露脐	短款	中款	长款	超长				
	袖长	吊带	无袖	短袖	中袖	长袖				
	领口	毛领	立领	V领	一字领/大开领	圆领	翻领	高领	有连衣帽	
	开口(多选)	单排纽扣	半排纽扣	双排纽扣	套头	拉链	半排拉链	敞开		
	版型	紧身	直筒	宽松	斗篷型	收腰				
	纹理(多选)	纯色	格子	圆点	碎花	横条纹	竖条纹	数字+字母	图案集中	图案重复
	材质(多选)	混纺	针织	丝绸/雪纺/纱	皮革	牛仔	纯棉	化纤	亚麻	呢绒
	属性(多选)	西装	运动服	冲锋衣	羽绒服	衬衫	毛衣	T恤	卫衣	大衣
		皮衣	夹克	马甲	皮草	连衣裙	旗袍			
<b>裤子</b>	<b>性别</b>	男	女	通用						
	长度	短裤	五分裤/中裤	七八分	长裤					
	腰部	高腰	低腰	普通						
	纹理(多选)	纯色	格子	圆点	碎花	横条纹	竖条纹	数字+字母	图案集中	图案重复
	材质(多选)	混纺	针织	丝绸/雪纺/纱	皮革	牛仔	纯棉	化纤	亚麻	呢绒
	版型	直筒	紧身	宽松						
	属性(多选)	西裤	工装裤	休闲裤	运动裤	热裤	裤袜	哈伦裤	喇叭裤	背带裤
<b>裙子</b>	<b>长度</b>	短裙	半裙	长裙						
	版型	包臀裙	A字裙							
	纹理(多选)	纯色	格子	圆点	碎花	横条纹	竖条纹	数字+字母	图案集中	图案重复
	材质(多选)	混纺	针织	丝绸/雪纺/纱	皮革	牛仔	纯棉	化纤	亚麻	呢绒
	褶皱	有	无							

图 11: 数据库完整标注特征列表