

刘雁兵,孔维力,刘晓蓉,等.基于多模态信息融合的卷烟销量预测方法[J].江西师范大学学报(自然科学版),2023,45(5):497-505.

LIU Yanbing, KONG Weili, LIU Xiaorong, et al. The multimodal-based method for tobacco sales forecast [J]. Journal of Jiangxi Normal University (Natural Science) 2023, 47(5) : 497-505.

文章编号: 1000-5862(2023) 05-0497-09

基于多模态信息融合的卷烟销量预测方法

刘雁兵¹, 孔维力¹, 刘晓蓉², 王义新², 汪伟飞³

(1.广西中烟工业有限责任公司, 广西南宁 530001; 2.广东烟草广州市有限公司, 广东广州 510510;

3.武汉人工智能研究院, 湖北武汉 430010)

摘要: 融合多模态信息的数据科学对智能营销至关重要. 该文提出了一种融合了视觉、自然语言和结构化数据的基于多模态信息的烟草销售预测方法. 首先, 引入扩散模型来生成高质量的香烟图像样本以进行数据扩增. 在柜台香烟识别阶段, 采用深度耦合网络和多元组排序损失来提高香烟识别率; 其次, 在销售预测中, 使用多模态信息作为输入, 包括柜台位置、采用双向文本编码的品牌表示以及相应价格; 最后, 通过预测算法得到香烟的预测销量. 通过全面综合的分析为营销提供了有价值的建议, 促进了多模态信息在卷烟营销科学上的应用.

关键词: 深度学习; 图像生成; 目标检测; 卷烟识别; BERT 模型

中图分类号: TP 391.4 **文献标志码:** A **DOI:** 10.16357/j.cnki.issn1000-5862.2023.05.09

0 引言

在烟草销售中,为实现智能营销^[1-3],数据科学^[4-6]越来越重要.调查产品信息和销售数据可以促进数据驱动的零售智能系统发展.香烟销售通常受到各种因素的影响,如价格、展示策略和品牌等.一般来说,香烟展示策略在营销中占据重要地位.在规则的形状和相似的大小中,香烟盒通常以统一的形式展示,一些畅销品有自己的热门位置.本文提出了一种基于多模态信息的销量预测方法,包括来自销售(POS)终端的信息和来自展示柜台的图像.

香烟的展示一直是其营销中的重要方面.穆建军^[7]创新地采用了3维展示方法,克服了小型商业区域的不便.刘薇^[8]探索了香烟的展示规则,以刺激消费者购买意愿.由此可以看出,香烟的展示在香烟营销中至关重要.通过使用智能零售终端来分析香烟品牌、展示方式和销售情况之间的关系.由于香烟

盒看起来形式相似,因此有必要构建一个细粒度对象识别模型.

本文提出了一种香烟销量预测方法,建立了销售信息和展示图像之间的联系.根据柜台上的展示位置、香烟品牌和相应价格预测每个品牌的月销量.在该方法中,采用扩散模型进行数据扩增;使用深度耦合对象识别网络和多元组排序损失获得陈列位置,以高效地区分不同类型的烟盒;使用中文双向编码器 BERT (bidirectional encoders representations from transformers) 获得品牌编码表示,将价格以结构化信息表示并输入添加线性回归模型以预测每月销量.

1 相关工作

1.1 相关背景

随着人工智能技术的不断发展,图像检测识别技术取得了长足的进步^[9],近年来在烟草行业的有

收稿日期: 2023-02-11

基金项目: 国家自然科学基金(62076235)资助项目.

作者简介: 刘雁兵(1978—),男,湖北洪湖人,博士研究生,主要从事计算机视觉和大数据技术应用的研究. E-mail:

wangweiwei@wair.ac.cn

关应用也不断增多,这集中体现在卷烟识别模型的迅猛发展。例如,肖雷雨等^[10]利用深度神经网络对烟梗形态进行检测识别,梁冬等^[11]提出了一种基于图像识别技术的卷烟零售数据采集方法,单宇翔等^[12]提出了基于深度学习的在复杂场景中的卷烟烟盒检测与识别方法,使香烟检测与识别的有关技术迅猛发展。

然而,在实际采集过程中,样本获取困难、样本分布不均衡等问题依然对模型的检测性能构成了不小的挑战。而传统的样本扩增方法主要包括随机裁剪、翻转等方式,并不能很好地对样本进行较大规模扩充,局限性较大。随着图像生成技术的不断发展,采用生成式模型扩充样本数据可以有效缓解此类问题。长期以来,生成式任务主要是使用生成对抗网络(GANs)进行的^[13-14]。最近,越来越多的工作使用扩散模型进行图像生成,其基本思想是通过迭代正向过程来系统地缓慢破坏数据分布的结构,并通过学习反向过程恢复数据的结构,并提出了一系列无条件或有条件引导的扩散生成模型。这些模型取得了极具竞争力的结果,甚至在许多图像生成任务中超过了基于 GAN 的模型^[15-16],成为目前最为流行和先进的生成式架构。

1.2 销量预测方法技术现状

香烟销售预测任务的主要方法包括时间序列法、机器学习法和深度学习法。在基于时间序列方法的香烟销售预测研究中,E.S.Gardner^[17]指出指数平滑预测在拟合短期到中期数据方面具有较高精度,可以有效利用历史上每个时期的数据。此外,多元回归方法在时间序列分析中也很常见。赵亚雄

等^[18]使用多元回归方法准确预测了物流需求。灰色预测模型 GM(1,1)也是时尚产品销售的常见预测方法。谢乃明等^[19]指出:当数据不足时,灰色预测模型仍然可以取得相对较好的结果。这些基于时间序列的预测方法在某些情况下存在局限性,不能被广泛使用。由于机器学习方法在回归预测任务中表现良好,因此已经有许多研究使用机器学习算法进行销售预测。孙晓静等^[20]分析了影响服装销售的各种因素,并发现服装销售趋势具有复杂的非线性特征。他们应用聚类算法对影响因素进行分类和分析,并使用 CART 决策树算法预测服装销售。与传统机器学习方法相比,深度学习方法在处理非结构化数据方面具有明显的优势。在深度学习中的深度神经网络可以将数据任务投影到高维空间进行处理,从而处理更复杂的预测任务。例如,在这项研究中,为了保留烟草品牌名称之间的相关信息,首先使用 BERT 获取品牌名称的嵌入,然后与 MLP(多层感知机)模型结合进行销售预测,进一步提高了预测精度。

2 基于多模态信息融合的卷烟销量预测方法

2.1 整体模型架构

该方法可以分为 3 个部分:第 1 部分主要集中在数据扩增上,使用条件扩散模型生成虚拟香烟的样本。第 2 部分主要为香烟盒识别网络,该网络预测了产品名称和显示位置。第 3 部分为销售预测,使用基于 BERT 的预测器来整合多模态信息并预测月销售额。整体框架如图 1 所示。

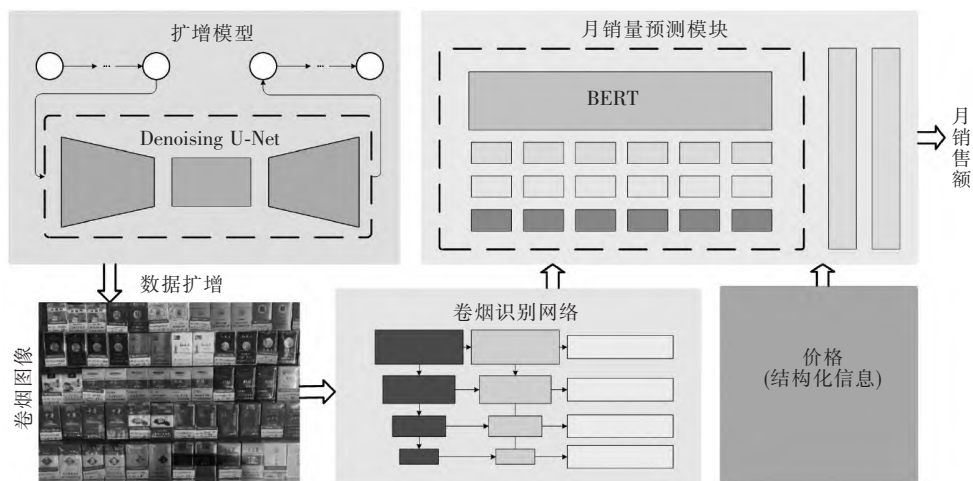


图 1 卷烟销量预测模型框架

2.2 基于扩散模型的卷烟样本生成

由于收集的香烟种类繁多,一些香烟类别的展示率较低,在收集的图像中覆盖这些类别的图像数量相对较少,因此在模型训练期间无法对该类别进行充分训练,这影响了模型检测效果。为了缓解这个问题,该文引入了基于扩散模型的虚拟香烟样本生成技术,对香烟样本进行数据增强,以提高模型的检测效果和增强模型的鲁棒性。卷烟图像生成网络如图2所示,在前向过程中,通过逐渐向 x_0 添加噪声来获得有噪声的香烟图像 x_t ,而在反向过程中,使用去噪网络逐渐生成干净的香烟图像,其中 x_1, x_2, \dots, x_t 是在扩散过程中产生的变量,与 x_0 具有相同的大小。

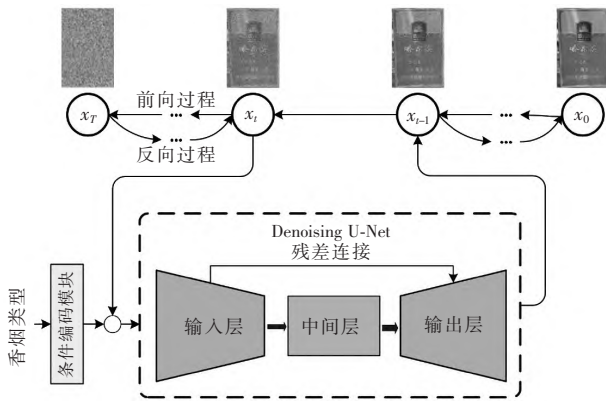


图2 卷烟图像生成网络

扩散模型的前向过程是一个 Markov 过程,可以根据以下概率分布逐渐向香烟数据 x_0 添加噪声,其中 $\beta_1, \beta_2, \dots, \beta_t$ 构成从0到1的方差时间表。最终,图像会趋向于高斯噪声 x_T 。

$$q(x_t | x_{t-1}) = N(x_t; \sqrt{1-\beta_t} x_{t-1}, \beta_t I),$$

在反向过程中,可以建立条件去噪网络 θ ,

其中 $\mu_\theta(x_t, t, c)$ 和 $\Sigma_\theta(x_t, t, c)$ 分别表示网络预测的均值和方差, c 表示香烟图像的类别作为条件输入。

$$p_\theta(x_{t-1} | x_t, c) = N(x_{t-1}; \mu_\theta(x_t, t, c), \Sigma_\theta(x_t, t, c)),$$

由于高斯过程的特性,可以通过给出时间戳 t

和 x_0 来直接获得 x_t 。令 $\alpha_t = 1 - \beta_t$, $\bar{\alpha}_t = \prod_{s=0}^t \alpha_s$, 则

$$q(x_t | x_0) = N(x_t; \sqrt{\bar{\alpha}_t} x_0, (1 - \bar{\alpha}_t) I).$$

损失函数 $L = L_{\text{MSE}} + \lambda L_{\text{VLB}}$ 结合了均方误差(MSE)损失和变分下限(VLB)损失,其中 λ 表示平衡这2种损失的超参数。MSE损失度量估计噪声与目标噪声值 ε 之间的差异,VLB损失优化了估计分

布与后验之间的KL散度:

$$\begin{aligned} L_{\text{MSE}} &= E_{t, x_0, \varepsilon} (\| \varepsilon - \varepsilon_\theta(x_t, t, c) \|^2), \varepsilon \sim N(0, I), \\ L_{\text{VLB}} &= D_{\text{KL}}(q(x_T | x_0) \| p(x_T, c)) + \\ &\sum_{t>1} \underbrace{D_{\text{KL}}(q(x_{t-1} | x_t, x_0) \| p_\theta(x_{t-1} | x_t, c))}_{L_{t-1}} - \\ &\sum_{t>1} \underbrace{\log p_\theta(x_0 | x_1, c)}_{L_0}. \end{aligned}$$

神经网络的输入包括有噪声的图像和香烟类别信息。类别信息通过类别编码块与网络特征融合。图像去噪网络使用深度U-Net网络作为基本架构,主要包括输入块、中间块和输出块。具体而言,输入块由17个残差块、6个注意力块和10个下采样块组成,中间块由2个残差块和1个注意力块组成,输出块由23个残差块、9个注意力块和10个上采样块组成。最后,输出添加噪声的均值和方差。

在采样阶段,使用无分类器引导技术^[16]来提高生成样本的准确性(见图3)。在采样阶段 t ,最终预测结果通过将输出与无条件引导融合得到,

$$\hat{\mu}_\theta(x_t, t, c) = \mu_\theta(x_t, t, \emptyset) + \lambda (\mu_\theta(x_t, t, c) - \mu_\theta(x_t, t, \emptyset)),$$

其中 s 表示引导尺度, \emptyset 表示空输入, c 表示香烟类别。

随后,通过下式计算 x_{t-1} :

$$x_{t-1} = (x_t - (1 - \alpha_t) \hat{\mu}_\theta(x_t, t, c)) / \sqrt{1 - \beta_t + \Sigma_\theta(x_t, t, c)^{1/2} n} / \sqrt{d_t}.$$

其中 n 是具有平均值0和方差1的正态分布随机噪声, $\Sigma_\theta(x_t, t, c)$ 是网络有条件引导下的噪声输出。

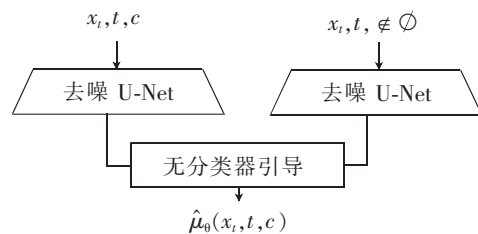


图3 无分类器引导技术

2.3 基于深度耦合网络的卷烟样本识别

在这个模块中,识别在柜台上展示的香烟盒。如图4所示,香烟盒定期在柜台上展示,并由工作人员拍摄图像。通过精细的模型,可以检测到相应产品的展示位置,其表示为 $p_{os} = (x, y, w, h)$,其中 (x, y) 表示香烟盒左下顶点的坐标, (w, h) 表示香烟盒的宽度和高度。通过将识别出的香烟与POS机器的销售

数据进行链接,可以全面了解产品概况.在处理数据之前,快速评估图像质量,并将一些不合格的图像(如过曝、过暗和运动模糊)进行过滤.



图4 卷烟图像示例

如前文所述,网络应该被精心设计以捕捉香烟盒上的细节.网络由带有注意机制的主干、特征金字塔网络(FPN)和一些耦合头组成(见图5).在主干网络中,设计了级联注意^[21]来捕捉显著特征.注意力块插入主干中,从粗到细自动学习对象的空间区域.FPN^[22]捕获多层次分层特征,并增强小目标的特征定位灵敏度.最后,耦合头集成多尺度特征,提取目标的局部组件、全局结构和上下文特征,形成目标特征的综合表达,并减少复杂光照引起的误识别.

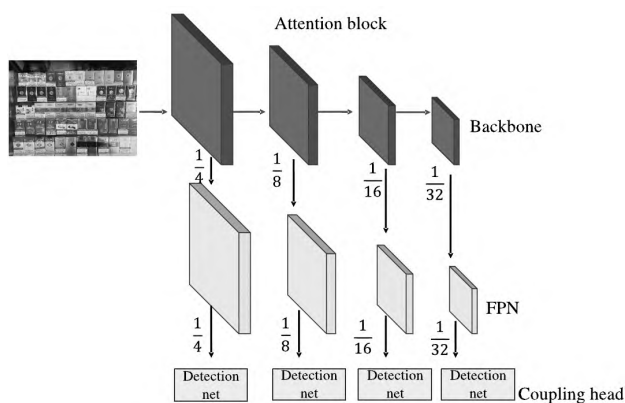


图5 深度耦合网络框架

除了精心设计的网络外,还利用多元组排序损失进行细粒度识别.通常,一个品牌有一系列产品,这些产品在外观上相似,但只有一些在文本或图案上的差异.这些小差异给细粒度识别带来了巨大的挑战.在这项工作中,设计的损失由一个目标图像3元组组成,其中 I_a 和 I_p 属于同一类别, I_n 来自不同的类别.

$$\varphi_1 = \|f(I_a) - f(I_p)\|_2, \quad (1)$$

$$\varphi_2 = \|f(I_a) - f(I_n)\|_2, \quad (2)$$

$$T(I_a, I_p, I_n) = \max\{\varphi_1 - \varphi_2 + m, 0\}. \quad (3)$$

如式(1)~式(3)所示,设计的损失降低在特征空间中 $f(I_a)$ 和 $f(I_p)$ 之间的欧氏距离,同时要求 $f(I_a)$ 和 $f(I_p)$ 之间的欧氏距离至少比前者大一个值 m .该损失可以尽可能地拉近来自同一产品的特征,并增强学习特征的区分能力.它保证了类内紧密性和类间区分性.

2.4 基于多模态信息的销量预测网络

探索不同位置的特定香烟品牌的销售情况可以帮助零售商做出更好的决策.首先分析展示位置和月销售额之间的联系,然后构建销售预测模型.一些零售商不喜欢使用POS机器,而是以非常传统的方式结账.这会在数据库中引入偏差或噪声.在处理数据之前,通过过滤那些每月离线超过10 d的零售商来选择高质量的POS机器零售商.

为了找到最受欢迎的品牌,绘制了一个香烟产品的热力图,例如,图6是一个地区平均之后的热力图.热力图是一个 5×12 的网格,图中的横纵坐标表示卷烟的位置,每一小格对应在卷烟陈列图片中某一种卷烟的规格,共呈现了60个烟盒的销量情况.从热力图中可以直观看出卷烟销量情况以及该规格在烟柜中陈列的位置,

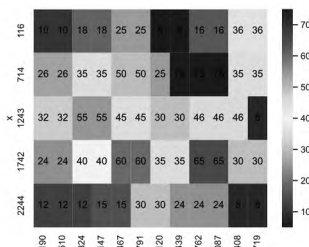


图6 香烟产品热力图

在这个阶段,考虑检测到的位置、产品价格和产品文本描述构建了一个销量预测模型.使用预训练的中文BERT做品牌描述.BERT是谷歌于2018年提出的一种预训练语言模型,从大量未标记的语料库中提取信息.网络结构由多个双向变压器编码器堆叠在一起组成.本文使用基于字符的BERT-base模型,它包含12个网络层、768个隐藏层、12个注意力头和1.1亿个模型参数.产品描述以字符嵌入和位置嵌入向量的形式进行编码.如图7所示,将展示位置特征 (x, y, w, h) 、产品价格和BERT特征连接起来,并将集成后的特征输入到销售预测器中.

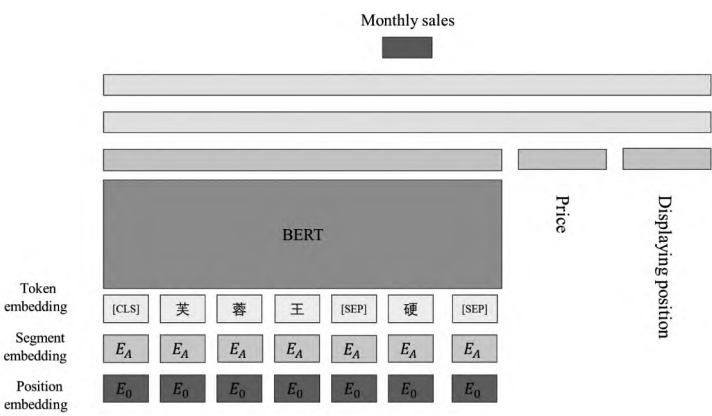


图 7 基于 Bert 的香烟预测网络

香烟预测模型采用的损失函数为 Huber 损失, 用于回归预测问题, 并具有一定的鲁棒性. 与均方误差 (mean squared error, MSE) 相比, Huber 损失对离群值的影响较小. 在 Huber 损失函数中, 对于误差较小的样本, 采用均方误差作为损失函数. 而对于误差较大的样本, 采用绝对误差作为损失函数. 本模型的损失函数定义如下:

$$L = \begin{cases} (\varepsilon - \varepsilon_\theta(B(N_{sk}), p_o, p_r))^2, & |\varepsilon - \varepsilon_\theta| \leq \delta, \\ \delta |\varepsilon - \varepsilon_\theta(B(N_{sk}), p_o, p_r)| - \delta^2/2, & |\varepsilon - \varepsilon_\theta| > \delta, \end{cases}$$

其中 ε 为真实值, ε_θ 为预测值, δ 为 Huber 损失函数的阈值参数, $B(N_{sk})$ 表示 BERT 的编码输出, N_{sk} 表示香烟名称, p_o, p_r 分别表示位置信息和价格信息. 当误差小于等于阈值 δ 时, 使用均方误差; 当误差大于阈值 δ 时, 使用绝对误差. 这样的设计可以在一定程度上平衡对异常值的处理, 提供了更好的鲁棒性.

3 实验数据与评价指标

3.1 实验数据

本文实验采用的实验样本为卷烟零售终端中采集到的前柜、背柜图像, 图像中包含店铺销售的不同品牌的卷烟小盒图像, 卷烟种类共计 207 类, 包含了国内 199 种卷烟和国外 8 种卷烟. 每类卷烟包含若干数量样本, 样本数量从数十张到数千张不等, 卷烟类别的数量分布图如图 8 所示.

3.2 实验方法

针对香烟识别模块的训练, 由于部分香烟的样本数量较少, 因此采用条件扩散模型生成虚拟烟盒, 以增强识别网络的性能, 扩充数据并缓解类别不均衡的问题, 最终确保每种卷烟的图片数量至少达到 500 幅. 柜台图像包含了真实的香烟陈列图像,

总共包括 1 200 个商店的 5 000 幅图像, 采用 4 500 幅图像进行训练并采用 500 幅图像进行检测.

针对香烟销量预测模块的训练, 采用识别网络获取得到香烟的名称、位置, 并结合价格和销量信息构建了 50 000 个数据样本, 并按照 7:3 的比例划分训练集和测试集.

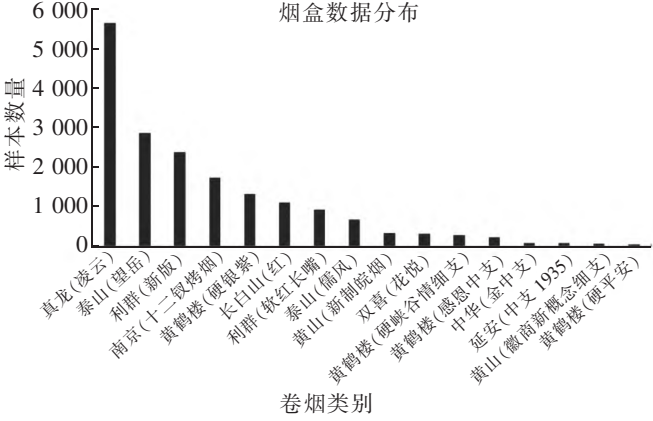


图 8 卷烟数量分布

3.3 实验评价指标

卷烟样本识别选用的评价指标为类别识别平均精确 (average precision, AP) 和全部类别均值 (mean average precision, mAP) 对烟盒识别模型进行评估.

针对卷烟销量预测模型, 本文选用平均绝对误差 (e_{MAE})、均方根误差 (e_{RMSE}) 和决定系数 (R_2) 作为评估模型性能的评价指标, 它们的计算公式分别为

$$e_{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|, \quad e_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2},$$
$$R_2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (\bar{y} - y_i)^2},$$

其中 \hat{y} 表示预测值, \bar{y} 表示平均值, y_i 表示真实值. e_{MAE} 、 e_{RMSE} 数值越低表明预测效果越好, R_2 数值越高表明预测效果越好.

4 实验结果

4.1 方案设计

有条件引导的扩散模型部分的扩散步数设置为1 000,在采样时的步数设置为250,无分类器引导的系数设置为2,学习率为0.000 1,训练轮数为100, batch 大小为8;卷烟识别网络的训练轮数为72,学习率设置为0.001;卷烟预测网络的学习率为0.01,训练轮数为30, batch 大小为64.上述训练采用的优化器均为Adam(adaptive moment estimation),既能适应稀疏梯度,又能缓解梯度震荡的问题,十分易用.本文的生成模型在4块 NVIDIA Tesla V100 显卡中进行,其他实验在4块 NVIDIA GeForce 1080 Ti 显卡中训练.

4.2 测试结果分析

4.2.1 数据扩增与卷烟识别效果 图9展示了部分虚拟卷烟样本的生成效果,由图9可见本模型可以生成较高质量的虚拟卷烟样本,所生成的样本满足对香烟样本进行扩充的需求.虽然在文字上生成的准确率还有待提高,但烟盒整体的生成效果已经显示了该品牌香烟的特征.

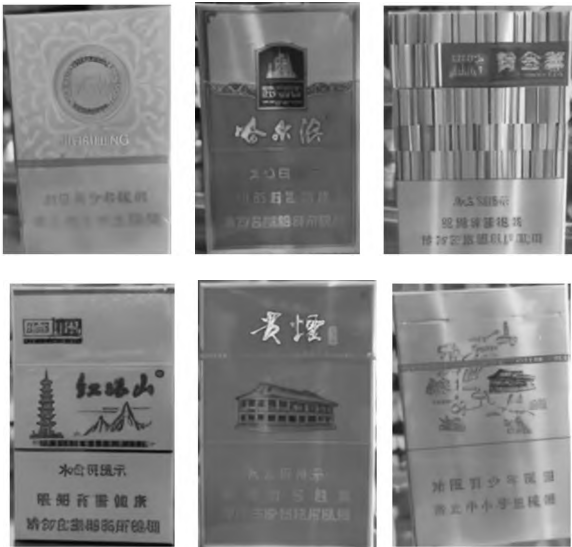


图9 部分虚拟卷烟样本示例

图10展示了部分香烟检测可视化结果,选取了有代表性的2种柜台进行展示.由图10可见:本模型可以较好地识别出不同香烟的名称、类别,并给出香烟的具体坐标位置信息.



图10 香烟检测可视化结果

表1展示了在采用不同模型下训练不包括(w/o)生成样本以及训练包括(w/i)生成样本的mAP指标情况.由此可见,采用了ResNet50+multi-group ranking loss训练策略的整体效果最好.在增加了生成样本后识别网络的性能有所提升.

表1 香烟样本识别性能测试

模型	mAP(↑)	
	w/o 虚拟样本	w/i 虚拟样本
ResNet50	0.858	0.932
ResNet50+coupling network	0.906	0.948
ResNet50+multi-group ranking loss	0.935	0.953

为了进一步评估虚拟样本增加对不同类型香烟识别性能的影响,表2展示了样本数量最多的5种类型香烟以及样本数量最少的5种类型香烟的识别准确率.由表2可见:由于样本数量本身较多,所以是否增加虚拟样本对样本数量较高的香烟类型识别精度的影响不大,但可以提升较少样本数量卷烟种类的识别精度.

4.2.2 卷烟销量预测效果 训练卷烟预测网络需要输入样本名称、位置信息、价格等信息,采用真实销量作为监督,训练的loss曲线如图11所示,横坐标表示训练轮数,纵坐标表示loss值.由图11可见,随着训练的深入,模型的loss不断下降,在训练达到400轮后,预测网络逐渐收敛,loss维持在相对较低的水平.

表 2 部分类别香烟识别指标

卷烟种类	AP(\square)	
	w/o 虚拟样本	w/i 虚拟样本
真龙(凌云)	0.968	0.970
泰山(望岳)	0.977	0.965
利群(新版)	0.982	0.984
南京(十二钗烤烟)	0.954	0.952
长白山(红)	0.961	0.972
黄鹤楼(感恩中支)	0.855	0.904
中华(金中支)	0.834	0.932
延安(中支 1935)	0.765	0.865
黄山(微商新概念细支)	0.786	0.897
黄鹤楼(硬平安)	0.744	0.886

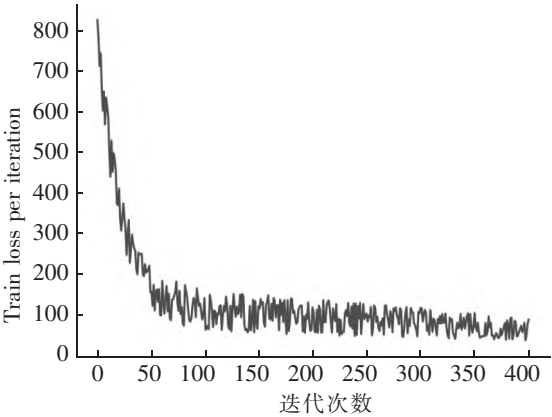


图 11 卷烟预测模型训练 loss 曲线

表 3 直观地展示了在测试集中随机选取的部分样本的预测销量与真实销量对比,样本为香烟名称,位置包括坐标及长宽(单位为 cm),展示了预测销量和真实销量的数值(单位为盒),由表 3 可见:其平均误差在 3 盒左右,最大误差不超过 5 盒.由此可知,本模型预测的销量与真实销量差异相对较小,可以较好地预测香烟的销量信息.

为了评估基于 BERT 模型的卷烟预测模型的性能,替换了其中的部分模块进行对比实验.表 4 展示了 e_{MAE} 、 e_{RMSE} 以及 R^2 的结果.按照卷烟销量对陈列位置的敏感程度划分为高敏感、中等敏感和低敏感 3 类,评价的模型包括采用 one-hot 向量以及 BERT 模型编码香烟信息以及采用支持向量机 SVM 和 MLP 作为分类网络进行预测的情况.根据评价结果,具有随着敏感程度的升高,模型的预测性能会有所提升.采用 BERT+MLP 策略的模型具有最为优异的性能,它反映了网络设计的合理性.

表 3 部分样本的预测销量与真实销量对比

样本		位置(x y w h)		预测销量 (盒/月)	真实销量 (盒/月)
双喜 (硬世纪经典)	坐标/cm	$x = 21.90$	$y = 1.10$	24	22
	长宽/cm	$w = 22.30$	$h = 16.50$		
芙蓉王(蓝)	坐标/cm	$x = 16.25$	$y = 14.74$	15	12
	长宽/cm	$w = 17.81$	$h = 16.33$		
	长宽/cm	$w = 10.19$	$h = 8.63$		
双喜 (硬经典 1906)	坐标/cm	$x = 23.26$	$y = 1.37$	68	77
	长宽/cm	$w = 23.73$	$h = 2.10$		
黄金叶 (天香细支)	坐标/cm	$x = 19.99$	$y = 9.30$	5	5
	长宽/cm	$w = 21.36$	$h = 11.50$		
双喜(红邮喜)	坐标/cm	$x = 20.61$	$y = 11.80$	8	5
	长宽/cm	$w = 22.87$	$h = 13.56$		

表 4 香烟销量预测网络指标测试

香烟类型	Metrics	$e_{MAE}(\downarrow)$	$e_{RMSE}(\downarrow)$	$R_2(\uparrow)$
高敏感	Onehot+SVM	12.41	25.40	0.70
	Onehot+MLP	11.22	23.35	0.74
	BERT+ SVM	10.50	22.32	0.78
	BERT+ MLP	6.95	15.54	0.84
中等敏感	Onehot+ MLP	24.50	32.53	0.68
	Onehot+ SVM	30.25	40.35	0.65
	BERT+ SVM	25.60	35.54	0.72
	BERT+ MLP	20.42	30.45	0.76
低敏感	Onehot+ MLP	30.20	42.25	0.64
	Onehot+ SVM	28.45	38.60	0.66
	BERT+ SVM	26.35	36.85	0.67
	BERT+ MLP	24.85	34.25	0.70
平均准确度	Onehot+ MLP	22.37	34.10	0.72
	Onehot+ SVM	23.31	34.94	0.75
	BERT+ SVM	20.82	32.24	0.78
	BERT+ MLP	17.41	27.94	0.80

5 结语

本文提出了一种基于多模态信息的烟草销量预测方法,包括数据扩增模块、香烟识别模块、销量预测模型.为了研究香烟盒的展示位置与月销量之间的隐含联系,首先引入了一种细粒度的香烟识别模型,该模型具有深度耦合网络和多元组排序损失;为了提升模型识别效果,引入扩散模型,并为样本量较少的香烟类别生成虚拟样本数据.然后,通过预测网络的结果,集成位置、品牌和价格作为输入,构建得到基于 BERT 的销售预测网络.该方法可以成为在烟草智能营销中有益的工具,为烟草行业的营销提供有益的参考.

6 参考文献

[1] KRAMOLI Š J , KOPE Č KOV Á M. Product placement: a smart marketing tool shifting a company to the next competitive level [J]. Journal of Competitiveness ,2013 , 5(4) : 98-114.

[2] DE ANDRADE M , HASTINGS G , ANGUS K. Promotion

of electronic cigarettes: tobacco marketing reinvented [EB/OL]. [2022-09-13]. <https://doi.org/10.1136/bmj.f7473>.

[3] SIMÕES D M ,BARBOSA B ,FILIPE S. Smart marketing with the internet of things [EB/OL]. [2022-09-19]. <https://www.igi-global.com/gateway/book/192038>.

[4] DHAR V. Data science and prediction [J]. Communications of the ACM ,2013 ,56(12) : 64-73.

[5] 蒋庆,陶峰,章玉,等.基于大数据分析的车辆安全检测系统研究[J].江西师范大学学报(自然科学版),2022 ,46(6) : 649-654.

[6] 苗永春,程艳.离群点检测方法及其在大数据时代下的改进方法研究[J].江西师范大学学报(自然科学版),2014 ,38(5) : 454-458 ,495.

[7] 穆建军.中小型烟草零售终端卷烟陈列方式初探[J].科技风,2012(16) : 192.

[8] 刘薇.卷烟陈列标准研究[J].商,2014(25) : 244-245 ,241.

[9] 何金宝,胡秋宝,付志超,等.基于 DCGAN 和改进 YOLOv5s 的桥梁表面缺陷检测识别[J].江西师范大学学报(自然科学版),2022 ,46(6) : 655-660.

[10] 肖雷雨,王澍,刘渊根,等.基于深度学习技术的烟梗形态分类与识别[J].烟草科技,2021 ,54(6) : 65-74.

- [11] 梁冬, 陈智斌, 农英雄, 等. 基于图像识别技术的卷烟零售数据采集方法 [J]. 中国烟草学报, 2021, 27(5): 81-89.
- [12] 单宇翔, 龙涛, 楼卫东, 等. 基于深度学习的复杂场景中卷烟烟盒检测与识别方法 [J]. 中国烟草学报, 2021, 27(5): 71-80.
- [13] ZHANG Han, XU Tao, LI Hongsheng, et al. StackGAN: text to photo-realistic image synthesis with stacked generative adversarial networks [EB/OL]. [2022-09-26]. <https://ieeexplore.ieee.org/document/8411144>.
- [14] XU Tao, ZHANG Pengchuan, HUANG Qiuyuan, et al. AttnGAN: Fine-grained text to image generation with attentional generative adversarial networks [EB/OL]. [2022-09-28]. <https://arxiv.org/pdf/1711.10485.pdf>.
- [15] DHARIWAL P, NICHOL A. Diffusion models beat gans on image synthesis [EB/OL]. [2022-09-23]. <https://arxiv.org/abs/2105.05233>.
- [16] JONATHAN H, SALIMANS T. Classifier-free diffusion guidance [EB/OL]. [2022-09-12]. <https://blog.csdn.net/NGUever15/article/details/128282683>.
- [17] GARDNER E S. Exponential smoothing: the state of the art [J]. Journal of Forecasting, 2006, 22(4): 637-666.
- [18] 赵亚雄, 唐立新. 基于粗糙集多元回归的湖北物流需求预测 [J]. 现代商贸工业, 2019, 40(2): 31-32.
- [19] 谢乃明, 刘思峰. 离散 GM(1,1) 模型与灰色预测模型建模机理 [J]. 系统工程理论与实践, 2005, 25(1): 93-99.
- [20] 孙晓静, 高慧, 陈云. 基于聚类分析和决策树算法的服装销售预测模型 [J]. 中国管理信息化, 2015, 18(9): 64-67.
- [21] 胡瑛, 罗银, 张瀚文, 等. 基于注意力机制的 3D 卷积神经网络孤立词手语识别 [J]. 湖南工程学院学报(自然科学报), 2022, 32(1): 55-60.
- [22] 吴娇, 王鹏, 乔梦雨, 等. 基于改进多级特征金字塔的陆战场目标检测算法 [J]. 计算机应用与软件, 2020, 37(10): 155-161.

The Multimodal-Based Method for Tobacco Sales Forecast

LIU Yanbing¹, KONG Weili¹, LIU Xiaorong², WANG Yixin², WANG Weifei³

(1. China Tobacco Guangxi Industrial Company Limited, Nanning Guangxi 530001, China;

2. Guangdong Tobacco Guangzhou Company Limited, Guangzhou Guangdong 510510, China;

3. Wuhan AI Research, Wuhan Hubei 430010, China)

Abstract: Data science with multimodal information is of great importance for intelligent marketing. In the paper, a tobacco sales prediction method based on multimodal information is proposed, including visual, natural language, and structured data. Firstly, a diffusion model is introduced to generate high-quality cigarette image samples for cigarette recognition. In the cigarette recognition stage, a deep coupled network and multiple sets of ranking losses are employed to improve cigarette recognition at the counter. Secondly, in sales prediction, cigarettes are represented with multimodal information, including counter location, brand representation encoded with Bidirectional Encoders Representations from Transformers (BERT), and corresponding prices. Finally, the monthly sales of cigarettes are provided. Through comprehensive analysis, valuable strategic recommendations are provided for marketing, promoting the multimodal-based application in tobacco science marketing.

Key words: deep learning; image generation; object detection; tobacco recognition; BERT

(责任编辑: 冉小晓)