

# 体育赛事赞助商广告品牌曝光率的视频检测与分析技术

中国传媒大学 元方 林书羽

**摘要：**随着深度学习和计算机视觉技术的发展，目标检测技术的速度越来越快，准确程度也越来越高，图片或视频中的品牌 LOGO 检测作为目标检测的一个应用方向，可以在营销和广告活动中发挥重要的作用。本文使用 YOLOv3 网络算法对一段男篮世界杯比赛花絮视频进行了 LOGO 检测，并分析了百岁山和耐克两个品牌在视频中的曝光情况。然后在此基础上探讨了 LOGO 检测在营销活动中的应用，包括计算品牌知名度，衡量赞助活动效果，获得品牌反馈和检测非法 LOGO 使用等四个方面。

**关键词：**市场营销 广告效果 深度学习 目标检测

**Abstract:** With the development of deep learning and computer vision technology, the prediction of object detection is getting more fast. The brand LOGO detection in pictures or videos is an application direction of target detection, which can play an important role in marketing and advertising activities. Here, the YOLOv3 network was used to test the LOGO of a men's basketball World Cup game, and analyzed the exposure of the two brands of Baisuishan and Nike in the video. Then, based on this, the application of LOGO detection in marketing activities is carried out, including calculating brand awareness, reducing the effect of sponsorship activities, obtaining brand feedback and detecting illegal LOGO usage.

**Keywords :** marketing; advertising effect; deep learning; object detection

在市场营销活动中，一个非常重要的主题是检测品牌的曝光程度。例如体育比赛的赞助商，比赛转播视频中品牌的曝光次数、位置和大小直接与受众的注意程度相关，这些可以作为赞助效果的衡量指标。但是，对这种曝光的定量测量一度是比较复杂的，因为一场体育比赛的时间通常较长，而一次赞助活动又包含多场比赛，通过传统人工编码的方式对上述信息进行标注，需要花费大

量的人力物力，还容易出现错误。除了比赛视频，社交网络也是品牌营销活动的主战场之一。据估计，社交网络中超过 80%的帖子包含图片或视频，这种趋势不仅仅体现在 Instagram 这种图片社交网络中，也体现在微博和 Twitter 这种原本以文本为主的社交网络中，因此，只对文本内容进行分析是不足以满足研究人员的需求的，这种图像识别检测的任务正是 AI 系统的强项。随着深度学习的快速发展，计算机视觉算法准确性得到了很大提升，最常见的应用就是人脸识别和 OCR 任务。而现在，通过一类被称为目标检测的算法，通过少量的标注训练集，我们就可以让计算机快速准确地从大量图片中自动寻找到品牌 LOGO，并识别其位置、大小和曝光时长。



图 1 LOGO 识别示例

## 一、基于深度学习的目标检测技术原理及发展

目标检测算法可以基于预先标注对象及类别的图片集合训练模型，从而识别图片中特定对象的位置。算法将图片作为输入，类别作为输出，同时在目标周围绘制一个边界框。最常用的算法基于卷积神经网络（CNN），它代表一类神经网络结构，由输入层、卷积层、池化层、全连接层等组成。通过使用卷积核对输入图片做卷积运算从而得到新的特征。目标检测比起一般分类的一个难点在于，一幅图片中并不一定只有一个边界框，因此输出层的长度是可变的，无法用标准的卷积网络和全连接层进行建模。为了解决这个问题，最直观的一种

方法是首先从图片中提取不同的候选区域，然后用神经网络对目标在每个区域中的存在进行分类。Ross Girshick 等在 2014 年提出了一种被称为 R-CNN 的方法，使用选择搜索从图片中提取 2000 个区域，然后将它们输入到卷积层中提取特征，最后使用支持向量机（SVM）预测区域中对象是否存在。这种方法对一个标准数据集的测试达到了 62.4% 的准确率，比传统算法高出近 20%。但是，R-CNN 的缺点是每幅图片都包含 2000 个候选区域，因此计算时间较长，并且选择搜索的过程是固定的，可能生成不太适合的区域。因此，研究者在此基础上又开发了 Fast R-CNN 算法。Fast R-CNN 将图片直接输入 CNN 以生成卷积特征图，从中确定候选区域并将其重塑为固定大小，这大大加快了目标检测的速度。但是，这种方法仍然没有完全规避选择搜索过程。任少卿等人由此又设计了 Faster R-CNN 网络，将候选区域的产生直接交给一个卷积神经网络，另一个卷积网络用于特征提取。比起 Fast R-CNN 的速度又得到了极大提升，在测试中，Fast R-CNN 识别目标需要用 2.3 秒，而 Faster R-CNN 只需要 0.2 秒，这使得实时的目标检测成为可能。

前面的几种目标检测算法都通过区域进行目标定位，网络中输入的不是完整的图像。而 YOLO 试图仅用一个神经网络完成目标检测任务，YOLO 的含义是 “You Only Look Once”。下图为 YOLO 的基本架构，它由 24 个卷积层和 2 个全连接层组成，卷积层用于抽取图像特征，全连接层用于预测输出概率和坐标。YOLO 将图片划分网格化的区域，每个区域只负责预测至多一个目标，然后修正目标区域的范围。YOLO 比其他物体检测算法快几个数量级，在它基础上改进的 YOLOv2 和 YOLOv3 也是目前最流行的目标检测算法。

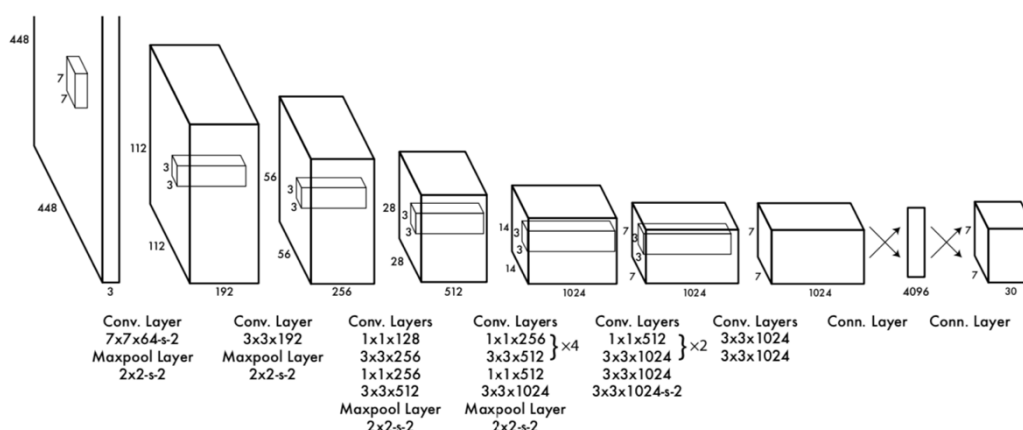


图 2 YOLO 网络架构

## 二、以男篮世界杯花絮视频 LOGO 检测为例的实验探究

在一场体育比赛中，研究人员通常会从多个方面评估品牌 LOGO 的曝光程度：

- (1) 曝光时长：LOGO 在屏幕上展示的时间越长，则效果可能越好。
- (2) 曝光位置：在比赛中，赞助商 LOGO 一般会出现在场边广告牌、地板或运动员服装上。通常如果 LOGO 出现在画面的中心位置，引起受众注意的可能性更大。另外，出现在优秀运动员衣服上的 LOGO 也可能增加品牌曝光的价值。
- (3) 曝光尺寸：毋庸置疑，覆盖更大空间的广告比较小空间的广告更有价值。
- (4) 曝光焦点：一般的，我们会认为如果同一个画面中没有其他品牌 LOGO 吸引受众的注意力时，品牌的曝光效果会更好。

这四个指标均可以通过 LOGO 检测结果分析得出。LOGO 检测的主要步骤包括 1) 处理视频，抽取帧图片并进行标注，2) 训练模型，3) 将模型应用于全部视频帧，输出结果，4) 汇总结果，绘制热图表示曝光程度。这里我们选择了一个比赛花絮片段用于实验，视频长度共 96 秒，每秒 25 帧。其中包含多个品牌 LOGO，我们选定百岁山矿泉水和耐克体育作为研究目标。

### 1. 标注训练图像

最常使用的视频流格式由每秒 24-26 帧组成，这意味着即使只有一分钟的视频，也可以分解成上千张图片。我们从这些图片中每秒抽取一张图片作为训练集，共 96 张图片，通过 labelImg 工具进行标注。



图 3 图片标注示意图

## 2. 训练并应用模型

LOGO 检测模型使用 YOLOv3 作为基本框架，在约 2000 次训练后达到最佳，Region AVG IOU 约为 0.68，该指标表示预测出的边框和实际标注的边框的交集除以并集，越接近 1 说明模型准确度越高。由于标注图片的数量有限，这个模型还有一些提升空间。得到模型后，我们将模型用于预测全部 2419 张图片的 LOGO 曝光情况，收集结果。前面的图 1 显示了预测标注框的样式。

## 3. 汇总结果并绘图

我们看到，作为男篮世界杯的主赞助商，百岁山一共出现在 2419 帧中的 1571 帧里，占总数的 64.9%，约为 63 秒，而耐克则一共出现了 926 帧，占总数的 38.1%，约为 37 秒。虽然耐克 LOGO 出现的时长没有百岁山 LOGO 多，但是如果计算 LOGO 曝光占总屏幕的面积占比，耐克的曝光面积为 6.78%，平均曝光面积为 1.58%，百岁山则只有 2.38%和 0.65%，在曝光尺寸上耐克更占优势。

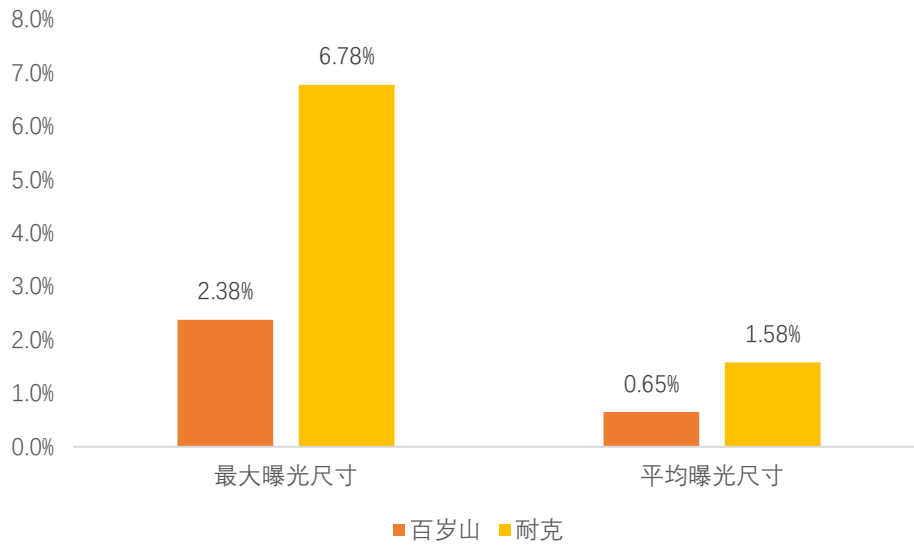


图 4 两个品牌的曝光尺寸

除了计算简单的百分比分析结果，我们还可以绘制热图，以查看整个视频中 LOGO 出现的位置。图中出现频率较高的部分（也就是高曝光区域）显示为红色，而很少出现的部分则显示为蓝色。百岁山和耐克 LOGO 出现的热图分别显示如下：

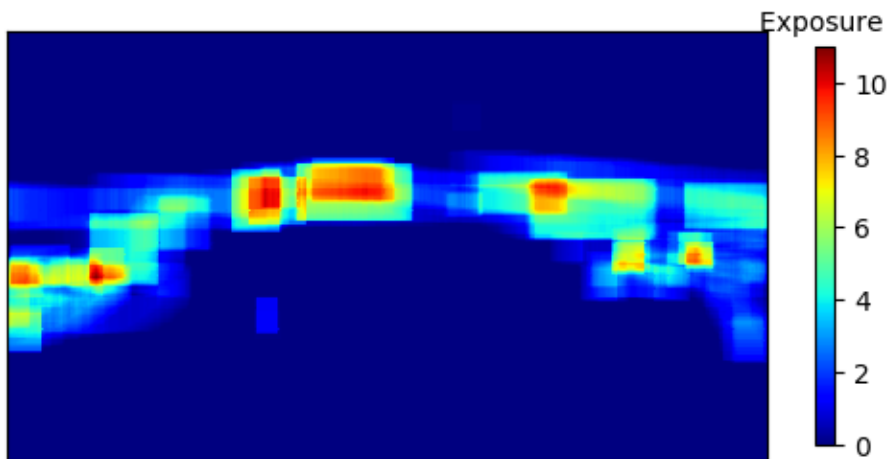


图 5 百岁山 LOGO 的曝光热图

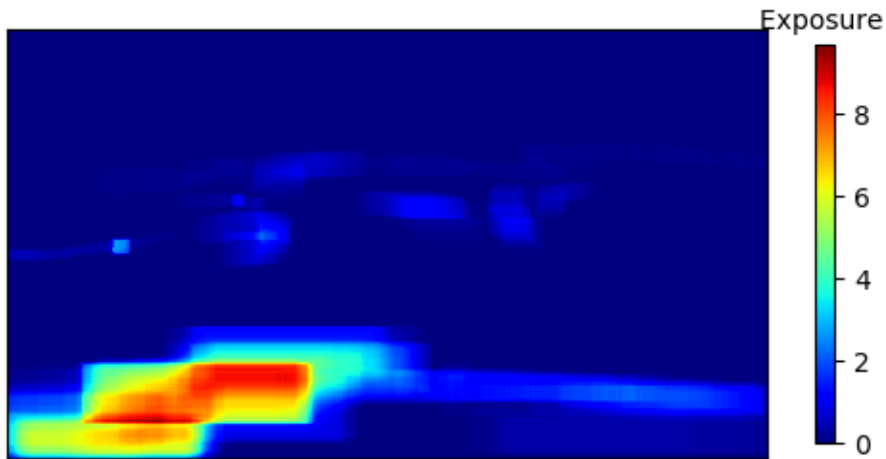


图 6 耐克 LOGO 的曝光热图

从图中可以很清晰地看出，百岁山的主要曝光区域集中在屏幕中间，略微呈现一个拱形，拱形的桥梁部分主要是场边的 LED 显示板，而两侧略微下降的部分则是篮架下的显示屏。耐克的主要曝光区域则在左下角，这里是球场地板的部分，另外也有少数曝光位置出现在场边。明显可以看到，场边显示板的曝光基本都位于屏幕中心，更容易进入受众的视线，但是地板上的 LOGO 曝光区域面积更大，这也与我们前述的计算一致。另外，百岁山在场边显示板的 LOGO 曝光时间更长，图 5 热图中最热的区域时长约为 11 秒，而耐克则为 9.7 秒。

也可以用一个折线面积图来查看两个品牌的 LOGO 同时出现的情况。下图横轴为秒数，纵轴为对应时间画面上的 LOGO 出现次数，可见，在该花絮视频刚开始时百岁山有一个很密集的曝光，而耐克则相对比较均衡，这样百岁山在视频开头就可能会得到比较高的注意。

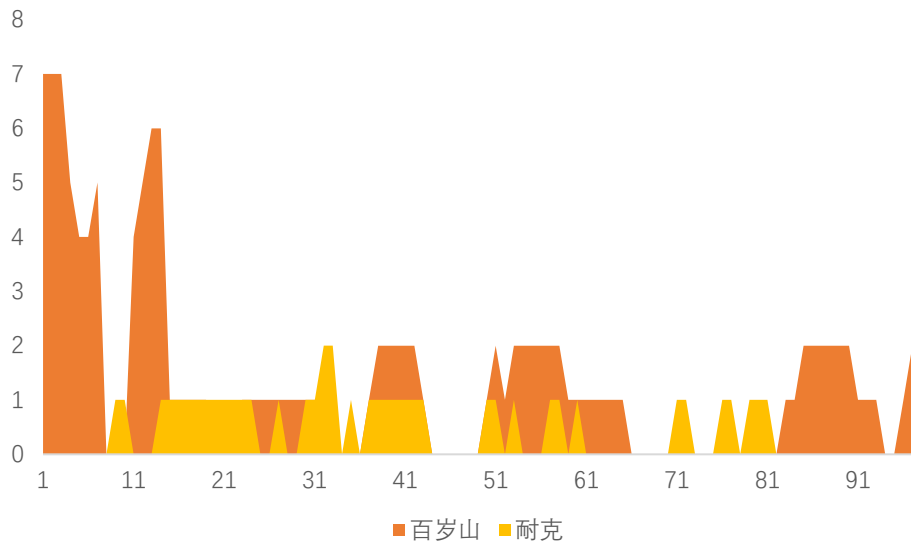


图 7 两个品牌 LOGO 随时间曝光次数

### 三、利用 LOGO 检测提升市场营销活动的有效性

在当下的媒体世界，受众无时无刻不在接触品牌的营销活动。通过监测相关的视觉内容，品牌厂商可以更好掌握广告品牌的曝光量，媒介与品牌的关联度，消费者表达的情感等等。品牌厂商还可以有效利用这种检测技术及时响应消费者的诉求。下面让我们总结一下 LOGO 检测可能应用于市场营销活动的几个方面。

#### 1. 计算互联网品牌知名度

一项来自哈佛大学的研究显示，大约 90% 的购买决策都是潜意识的，因此品牌始终保持足够的知名度是有意义的。根据 Hootsuite 的报告，过去两年中，Facebook 用户在广告上的参与度增加了两倍，到目前为止，超过 60% 的用户在 Instagram 上发现品牌和商品。社交媒体看起来就像是互联网时代的购物中心，因此，网络社区中品牌的知名度越高，就更可能影响消费者的购买决策，尤其是拥有众多追随者的意见领袖群体。但是，很多时候，品牌内容并不会出现在发布内容的文本和标签中，而是通过图片或视频展示，这时目标检测技术就可以派上用场，通过把图片或视频中的品牌曝光纳入计算体系，营销人员可以更准确地计算品牌知名度。

#### 2. 衡量赞助活动效果

成为大型活动的赞助商是迅速提高品牌曝光度的有效方式，例如大型体育



赛事或音乐节。但是，如何量化地衡量这些活动的效果是一个较难实现的目标。通过 LOGO 检测，研究人员可以计算电视节目中的品牌曝光度，也可以更全面地得到社交网络中的讨论热度，进而得到对赞助活动效果评估的量化指标。同时，由于这些活动的赞助费用通常很昂贵，有了这种量化的效果评估方法，计算投资回报率也成为可能，这样就可以更好地帮助营销人员，指导赞助活动。

### 3. 获得品牌反馈

受众反馈对营销人员来说也是非常重要的信息，尤其是其中的情感倾向。正面的反馈表示品牌的发展路径是正确的，而过多的负面反馈则表明某些方面需要得到调整。通过将图像检测和自然语言处理技术结合起来也可以得到更完善的品牌情感印象，也有一些更高级的算法可以直接从图片中评估视觉内容的倾向性。

此外，检测到 LOGO 后，研究人员也可以检索其他的数据，例如在主题标签、地理位置等，这可以帮助他们建立消费地图或分析出品牌联想。例如，一项研究发现苹果（Apple）在与猫出现在同一张图片上的品牌中排名第一，而阿迪达斯则在与狗的排名中名列前茅。这可能是由于猫喜欢卧在笔记本电脑上，宠物主人认为这样的画面很有趣，因此会拍照分享。而狗主人则可能经常穿着运动鞋与狗一起散步，这使得它们更容易同时出现在镜头中。这些研究可以使我们获得对品牌的新见解。

### 4. 检测非法 LOGO 使用

当品牌发展到一定知名度时，对品牌形象的保护就会变得尤为重要。这种保护包含了对品牌 LOGO 的滥用及非法使用的检测。通过 LOGO 检测工具，可以跟踪品牌 LOGO 及其变体的使用，也可以通过相似程度的检测来发现是否有“山寨”品牌的出现。在社交媒体上也有可能出现假冒伪造产品的照片，LOGO 检测有助于发现这类问题，进一步实现在线品牌形象的保护。

## 四、结论与讨论

本文基于计算机技术发展带来的图像和视频检测技术，针对产品品牌检测应用技术原理和发展进行了总结，以男篮世界杯花絮视频 LOGO 检测为例进行了实验探究。结果表明，计算机程序化图像检测技术在一定程度上可以替代传统的人工监播手段，已经取得了较高的准确度，并拓展丰富了量化的数据指标，为定量研究品牌传播效果提供了新思路和新方法，具有很大的价值。将计算机

发展的新技术，应用于传统的市场研究方法体系中，具有一定的创新性。

本文针对男篮世界杯花絮视频 LOGO 检测的研究是通过迁移学习将 YOLOv3 预训练模型用于 LOGO 检测任务的一次尝试，效果目前看来尚令人满意，但在丰富标注素材、场景、目标的基础上，其准确度仍有提升空间。将其用于商业使用，需标注大量的训练用素材库，或利用一些已有的标注品牌 LOGO 库，以提高模型广度，并根据检测目标适当调整训练素材，进一步提高模型深度和准确性。本文仅就品牌在视频中出现的位置、大小、时长等客观信息进行了讨论，品牌与视频图像中其它内容的关联性 etc 等主观性强的信息表达，可以作为进一步研究讨论的目标。

## 参考文献:

- [1] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition 2014 (pp. 580–587).
- [2] Girshick R. Fast r-cnn. In Proceedings of the IEEE international conference on computer vision 2015 (pp. 1440–1448).
- [3] Ren S, He K, Girshick R, Sun J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems 2015 (pp. 91–99).
- [4] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 779–788).
- [5] 周俊宇, 赵艳明. 卷积神经网络在图像分类和目标检测应用综述. 计算机工程与应用. 2017;53(13):34–41.
- [6] Coulter RA, Zaltman G, Coulter KS. Interpreting consumer perceptions of advertising: An application of the Zaltman Metaphor Elicitation Technique. Journal of advertising. 2001 Dec 1;30(4):1–21.
- [7] Cooper P. Facebook Stats That Matter to Marketers in 2019. <https://blog.hootsuite.com/facebook-statistics/>
- [8] Elena Vinokurtseva. What Is Logo Recognition? 7 Ways to Use It

for Your Marketing Campaigns. <https://learn.g2.com/logo-recognition>