



西北工业大学

NORTHWESTERN POLYTECHNICAL UNIVERSITY

研究生课程考试答题册

得 分：

学 号 _____

姓 名 _____

考试课目 电子商务技术

考试日期 _____

西北工业大学研究生院

西北工业大学研究生院

学 位 研 究 生 课 程 考 试 试 题

考试科目： 电子商务技术

课程编号：

考核形式： 课程论文

考试时间：

开课学期： 2026 年春至夏学期

任课教师：

说 明： 所有答案必须写在答题册上, 否则无效。

共 1 页 第 1 页

写一篇研究论文，要求如下：

- (1) 题目自拟；
- (2) 研究内容：针对你的研究领域，以电子商务为应用背景，研究 AI、LLM、知识表示、Agent 等 ICT 技术的应用。
- (3) 要求：论文有对领域问题的举例或分析、有对相关技术或者方法的归纳总结和比较分析，有应用案例。分析清楚，论述透彻，结构严谨，参考文献应是近年的高水平期刊的文章或者来自知名的研究机构的解决方案。书写格式规范，图文并茂，字数在 6000 字左右。

深度学习驱动的计算机视觉在电子商务商品识别与虚拟试穿中的应用研究

【摘要】随着电子商务向智能化、体验化快速转型，如何高效完成商品数字化识别、构建趋近线下的沉浸式购物体验，成为行业突破运营成本瓶颈、提升用户转化的核心命题。深度学习技术的成熟，推动计算机视觉从“被动感知”升级为“主动理解”，为电商场景的核心痛点提供了可落地的技术路径。本文以商品识别、虚拟试穿两大核心应用方向为研究对象，首先系统梳理二者在电商场景下的业务价值与技术难点；其次从技术架构演进、核心方案对比、场景化优化策略维度，分别归纳分析两类任务的主流深度学习技术体系；再结合头部电商平台的落地案例，补充技术与业务的融合逻辑；最后从技术协同视角出发，总结现有技术体系的共性瓶颈，并展望多模态、轻量化、大模型融合的未来发展趋势。研究表明，扩散模型、Transformer 架构、多模态融合作为当前技术升级的核心方向，已成为驱动电商运营效率、用户体验双向增长的关键基础设施。

【关键词】电子商务；计算机视觉；深度学习；商品识别；虚拟试穿；多模态融合

1 引言

近年来中国电商行业持续向场景化、体验化升级，行业内的核心矛盾也从“商品供给是否充足”转向“线上消费体验是否能匹配线下场景”。其中最突出的表现是，服饰、快消品等核心类目线上销售占比已突破 50%，这意味着超过半数的消费决策需要依托纯线上视觉信息完成；但与此同时，这类商品的退货率高达 30%-60%，远高于线下零售场景——其根源正是传统静态图片、文字描述的商品信息供给模式，无法支撑消费者完整、准确地获取商品感知。具体而言，这一基础信息缺口在实际业务端主要表现为两大核心痛点：

其一，对平台与商家而言，商品数字化管理的成本居高不下。电商平台的 SKU 规模动辄以千万级计算，新商品上架、库存盘点、效期排查等环节，都需要依赖

视觉技术完成商品识别、定位与特征提取。早期的人工标注、传统条形码识别等方案，要么是人力成本极高，要么是对复杂场景的识别能力不足，无法支撑大规模、高频率的商品数字化运营，这也成为限制电商运营效率提升的关键瓶颈。

其二，对消费者而言，“图不符实”已成为消费决策的核心障碍。在收到商品前，用户无法通过线上素材完整感知服饰的上身效果、面料质感、版型合身度，这是导致高退货率、低转化率的直接原因。行业研究数据显示，超过半数的消费者对线下试衣间的体验并不满意，而线上试穿的便捷性与私密性，恰好匹配了这类用户的潜在需求。但与此同时，全球时尚电商的平均退货率仍维持在 30% 以上，部分商家甚至突破 40%，这背后是用户对“真实试穿体验”的刚性需求，与现有技术能力不足之间的核心矛盾。

作为计算机视觉领域的突破性技术，深度学习的出现恰好为上述痛点提供了标准化的解决路径。与传统方法依赖人工设计特征的逻辑不同，深度学习通过端到端的多层神经网络结构，自动学习、提取商品的多层次视觉特征——从基础的颜色、纹理、形状等浅层特征，到领口设计、裙摆样式、logo 细节、五金配件等深层结构特征，甚至能捕捉不同光照、角度、包装下的差异化特征，从而实现更精准、更鲁棒的识别效果。在这一技术逻辑下，基于深度学习的计算机视觉技术在电商场景中的应用，逐渐分化为两条清晰的主线：

一是商品识别技术，作为电商数字化运营的基础支撑，覆盖商品分类、陈列状态巡检、自助结算、库存动态盘点等核心环节，目标是实现商品从生产端到消费端的全流程数字化管理；

二是虚拟试穿技术，作为连接商品与消费者体验的核心枢纽，将服饰商品的平面图像转化为贴合用户身材特征、能呈现真实上身效果的可视化内容，目标是补齐线上购物的体验短板。

当前，电商行业的技术探索已进入“场景化适配”的加速期：头部平台纷纷推出自研落地项目，行业商业化验证标准逐步明确，学术研究与工业落地之间的边界快速收窄。基于此，本文将结合近年行业实践案例与技术数据，深入分析商品识别与虚拟试穿的技术架构、适配逻辑、落地效能及演进路径，系统剖析深度学习技术在电商场景下的应用逻辑与发展趋势。

2 电子商务场景下的问题分析

在深入技术分析前，需先明确两类应用的业务边界、业务价值与技术难点——正是这些场景下的实际约束，决定了深度学习技术的选型标准、优化方向与落地策略。

2.1 商品识别的业务价值与技术难点

商品识别是指通过计算机视觉技术，自动检测、识别商品的类别、状态等核心信息的过程。这一技术是电商数字化运营的基础支撑，并非独立的业务环节，而是支撑上游业务高效运转的技术前提——其业务价值体现在三个核心维度：

一是运营成本优化：在智能货架、无人仓配、线下门店等全渠道场景下，机器视觉技术能替代人工完成商品分类、盘点、巡检等高频重复作业，显著降低人力成本。例如在商品上架环节，视觉分类系统通过商品外观特征与条码的双重核验，能将入库效率提升 40%，同时将人工核验环节的出错率降至 0.3% 以下。

二是消费体验升级：在自助结算、视觉搜索、商品推荐等面向用户的场景下，精准的商品识别能缩短用户的操作链路。最典型的是 AmazonGo 的“Just Walk Out”技术，用户完成商品挑选后无需主动扫码，系统就能自动识别商品并完成结算，极大优化了线下即时消费场景的体验。

三是数字化决策支撑：商品识别采集的全链路数据，是平台、商家进行数据分析、智能决策的基础依据。例如通过识别商品的陈列状态、用户拿起放回的行为数据，商家可以优化陈列位置、调整补货策略；某超市在应用相关技术后，将畅销商品的缺货率从 8% 降至 2%，同时将库存周转率提升了 15%。

从技术实现视角看，商品识别任务的核心逻辑并不复杂：通过摄像头采集商品图像，依次进行图像预处理、多层特征提取、目标匹配分析，最终完成商品类别、状态的判定。但电商场景的非理想性，对技术落地提出了远高于实验室场景的环境约束——这一技术从实验室走向真实业务场景，必须攻克四大典型技术瓶颈，也是近年商品识别技术优化的核心方向：

1. 类内差异小、类间差异大的细粒度识别难点：同类别商品的包装、设计高度相似，比如不同口味的饮料、不同功效的洗护用品，仅在标签文字布局、LOGO 色彩渐变、包装尺寸比例等细节上存在差异；但现有公开数据集的训练数据大多缺少这类差异化的细节标注，导致模型容易出现特征提取混淆，进而造成识别结果的偏差。

2. 陈列场景的遮挡、变形、角度非理想性难点：商品在实际场景中往往是堆叠放置、多层陈列，容易出现遮挡、反光、透视变形等问题；单一角度的摄像头采集到的图像信息往往存在局限性，无法完整呈现商品的核心特征，这直接提升了检测、识别的难度。

3. 实时性与识别准确率的平衡难点：电商场景的业务流量具有突发性、集中性特征，比如大促期间的自助结算环节，需要在毫秒级时间内完成商品识别结果的返回；而高精度模型的计算资源消耗更高，推理耗时更长，这就要求技术方案必须在保证准确率的前提下，兼顾实时响应效率。

4. 复杂环境干扰下的鲁棒性难点：电商场景的光线条件、拍摄设备存在不可控性：超市、仓库的现场照明光线强度分布不均匀，部分区域存在强烈的镜面反射，不同摄像头的成像参数、分辨率差异较大；这些无关变量会直接干扰特征提取的精度，要求模型对这类环境干扰具备鲁棒性。

2.2 虚拟试穿的业务价值与技术难点

虚拟试穿是补足电商体验短板、驱动业务增长的关键核心技术——其业务价值更贴近用户的消费决策，体现在以下三个层面：

其一，降低退货率，优化运营成本：这是虚拟试穿最直接的业务价值。消费者通过虚拟试穿提前掌握商品的上身效果，能有效降低因版型不合、材质效果预期偏差导致的退货率；对服饰电商而言，退货带来的逆向物流成本、商品损耗成本占运营成本的比例极高，虚拟试穿技术的落地能直接削减这部分无效支出。

其二，提升用户参与度，激活潜在消费需求：从用户行为链路看，虚拟试穿是将“平面商品素材”转化为“个性化穿搭参考”的关键节点，能将用户的被动浏览行为，转化为主动探索的交互行为。用户在找到贴合自身身型、风格偏好的穿

搭方案后，更容易形成明确的消费决策，进而提升商品转化率。

其三，支撑新业务模式的探索：虚拟试穿是构建全渠道数字化购物环境的基础支撑，能将线上、线下的消费场景通过体验链路打通——比如结合 AR 技术，用户在实体店可以通过手机查看更多尺码、颜色的试穿效果，无需在试衣间频繁更换；这一技术还能支撑“个性化定制”这类新业务模式的拓展，将消费需求与生产设计链路直接打通。

与商品识别聚焦“检测精度”的核心要求不同，虚拟试穿的技术难点更关注“生成效果的精准度”——需要完美匹配商品特征、用户特征与场景化体验要求，这对技术方案的逻辑完整性提出了极高要求。从用户的核心诉求视角看，虚拟试穿技术的落地需要逐层攻克三重技术难关，这也是制约该技术发展的核心瓶颈：

基础层难关：服饰细节的高保真还原：算法需要完整保留服饰的视觉特征——包括颜色、纹理、logo 等基础特征，以及面料质感、褶皱、垂坠感等动态特征。传统方案往往存在“纹理拉伸变形”“特征丢失明显”的缺陷，无法让用户获得足够真实的视觉感知；更关键的是，生成效果必须和用户的真实身材、面部特征自然融合，不能出现明显的 AI 合成痕迹。

核心层难关：合身度的精准模拟：这是当前虚拟试穿技术落地的最大瓶颈，也是用户最核心的真实需求。即使是同一款服饰，不同身材、体型的用户穿上后的实际效果也有明显差异；但电商平台缺少“小体型穿大码”这类非合身场景的真实数据，难以训练出精准适配身型的模型；再加上不同品牌的尺码标准不统一、用户对合身度存在个性化偏好，导致这一问题的技术复杂度远高于基础效果还原。

突破层难关：全场景的智能适配能力：用户的消费决策往往基于完整的场景化穿搭需求，而非单件商品的孤立效果；这就要求技术方案不仅能实现单件服饰的试穿，还要具备多件搭配的协同试穿能力，让模型能够结合用户的体型、风格偏好，输出适配不同场景的完整穿搭方案。

3 深度学习在商品识别中的技术归纳与分析

商品识别的技术架构经历了从“传统人工特征提取”到“深度学习自动特征提取”的完整代际演进，当前的主流方案是针对不同场景的技术约束，适配不同

的深度学习模型——没有任何一种模型能覆盖所有场景，核心逻辑是“场景化的精度与性能平衡方案”。

商品识别的核心任务是“定位商品位置、判断商品类别”，这一技术的演进路径，完全是对场景约束的适配升级。从技术架构的差异维度，可将其发展历程划分为两个典型阶段：

1. 传统机器学习阶段（2014 年以前）：这一阶段的技术方案逻辑是“人工设计特征+分类器拟合”，核心是依靠算法工程师人工设计的特征提取规则，比如 SIFT、HOG 等特征算子，来提取商品的边缘、纹理等基础视觉特征；再将这些特征输入支持向量机（SVM）等分类器，完成最后的类别判定。这类方案的优势是技术成熟度高，在背景简单、商品包装差异明显、光照条件理想的基础场景下基本可用；但面对相似包装、多视角陈列等复杂场景，特征提取的鲁棒性不足，容易出现识别准确率大幅下降的问题。

2. 深度学习阶段（2014 年至今）：这一阶段的技术方案逻辑是“端到端自动特征提取”，不再依赖人工设计的特征算子，而是通过多层卷积神经网络结构，自动从海量图像数据中学习、提取商品的浅层、中层、深层组合特征，再完成目标的检测、识别。这类方案的特征表达能力更强，能有效应对更复杂的实际应用场景——无论是识别精度、鲁棒性，还是场景适应能力，均远超传统机器学习方案。

在深度学习技术体系下，根据技术架构的差异，可将商品识别的核心技术方案分为三类，其特性、适配场景均有明确差异，核心是对不同场景技术约束的针对性适配：

1. 一阶段检测器：以 YOLO 系列、SSD 模型为代表，这类算法的核心特点是“单次前向传播完成计算”，将输入图像的特征提取、候选区域生成、边界框回归、类别置信度预测，统一整合在同一阶段完成；在保证较高检测精度的前提下，实现了极快的推理速度，更适配实时性要求高的场景。

2. 两阶段检测器：以 Faster R-CNN、Mask R-CNN 模型为代表，这类算法将检测过程拆分为两个串行阶段：第一阶段通过区域提议网络（RPN），生成图像中可能包含商品的候选区域；第二阶段对这些候选区域进行精细的特征提取、分类与边界框回归。拆分处理的方式能有效提升检测精度，但会增加计算资源消耗，

推理速度相对较慢，更适配对精度要求严苛、但实时性要求相对宽松的场景。

3. 基于 Transformer 的检测方案：以 DETR、ViT 这类架构为代表，这类方案的核心创新点，是将 Transformer 的全局注意力机制引入商品识别任务——与卷积神经网络只能提取局部特征的方式不同，这类架构可以在特征提取阶段，直接建模商品局部特征之间的全局依赖关系，进一步提升识别的鲁棒性，更适配细粒度、高遮挡场景的识别任务。

4 深度学习在虚拟试穿中的技术归纳与分析

虚拟试穿的技术目标是“将精准匹配的服饰特征叠加到用户图像上，生成满足真实感要求的合成结果”——其技术演进逻辑比商品识别更依赖生成式 AI 的进步，从早期的“几何变形+区域融合”思路，转向了“端到端生成式建模”方案。

虚拟试穿的核心技术难点，是在保留服饰自身属性特征的基础上，实现其不同人体姿态、身形、光照环境下的自然适配，还原真实上身效果——这一任务的技术复杂度远高于单纯的商品识别。从技术架构的维度，其演进路线可以划分为三个典型阶段：

1. 第一阶段（2017 年以前）：以“3D 建模 + 物理仿真”为核心逻辑。这一阶段的技术方案主要依赖专业级 3D 建模设备，精准捕捉人体的关键尺寸、体态特征，再通过物理引擎模拟服饰的面料特性，实现相对精准的穿戴效果。这类方案的优势是贴合度的精准度高，但缺点是设备成本极高，且流程的技术复杂度极高——不仅需要用户上传多张高清照片完成 3D 人体建模，还需要对服饰的多个维度数据进行同步采集，最终的合成效果质量高度依赖前期建模的精准度，无法面向普通消费者大规模普及。

2. 第二阶段（2017-2022 年）：以“生成对抗网络（GAN）”为核心技术逻辑。这一阶段的技术方案主要是通过 GAN 的生成器、判别器结构，完成试穿效果图的生成，核心逻辑是将服饰的几何变形与纹理生成分离为两个串行环节。这类方案的落地门槛显著低于 3D 方案，但仍存在明显的技术局限性：一是生成效果的分辨率上限较低，无法实现高清级别的输出效果；二是对复杂场景的细节表现力不足，无法真实模拟丝绸、皮革、薄纱等特殊面料的质感；三是在处理多维度、大角度

的姿态变化时，容易出现服装错位、纹理变形、细节丢失等问题。

3. 第三阶段（2023 年至今）：以“扩散模型（Diffusion Model）”为核心技术逻辑。这一阶段的技术方案能实现更逼真的视觉效果，更好地保留服装的细节纹理和人体的姿态特征；更关键的是，这类方案的工程落地门槛显著降低，更适配电商场景下的用户使用习惯。因此，扩散模型自 2023 年起，逐渐取代 GAN 成为虚拟试穿技术的主流方案；而基于扩散模型的升级架构，也成为了行业技术探索的核心方向。

在扩散模型成为主流技术方案后，行业内基于基础架构的优化方向差异，将主流虚拟试穿技术方案进一步细分为三类，其适配场景、效果表现均有明确差异：

1. 基于 U-Net 架构的扩散模型方案：这是早期扩散模型的典型实现方式，U-Net 模型作为扩散模型的核心生成器，通过编码器 - 解码器的对称架构，完成语义信息的特征提取与分辨率恢复，实现较高精度的试穿效果生成。

2. 基于 DiT（Diffusion Transformer）架构的方案：这是 2024 年后的主流升级方向，核心是将扩散模型的主体生成架构，从 U-Net 替换为视觉 Transformer（ViT）模块，利用 Transformer 架构的全局自注意力机制，提升生成结果的语义对齐精度；这一架构升级的逻辑，与同期 Sora 等生成式 AI 模型的技术演进趋势完全同步。

3. 基于 Mask Free 架构的方案：这是当前技术的重点优化方向，核心是引入视觉语言模型（VLM）进行多模态特征的统一表征，不再需要用户手动标记服装区域或进行复杂的人体解析，直接通过端到端的模型生成试穿效果。

5 典型行业应用案例分析

基于深度学习的计算机视觉技术，并非停留在学术研究阶段的技术方案，而是已经经过头部电商平台的商业落地验证，形成了成熟的业务闭环。本文选取国内外头部电商平台的落地案例，从技术选型、业务场景、落地效能、技术优化逻辑四个维度进行交叉对比，分析技术与实际业务场景的适配逻辑，以及技术落地的真实效果。

5.1 商品识别行业应用案例

商品识别的技术落地效果，往往作为基础能力间接地体现业务价值 —— 但从部分公开的项目数据中，仍能直观感受到技术优化对业务增长的驱动作用。本文选取了三个行业内具有代表性的落地案例，从技术选型、场景适配及实际业务效果的维度，进行详细的对比分析：

Amazon Go（美国）：作为“无人零售”概念的头部落地项目，其技术方案的核心架构，是计算机视觉技术与多模态传感器数据的深度协同融合——在店内部署了数百个高清深度摄像头，结合压力传感器、红外传感器等多种感知设备，覆盖从商品上架到用户选购、结算的全流程识别场景。在算法层面，以 YOLO 系列的单阶段检测器为基础骨干网络，结合多视角融合、多模态感知的优化技术，协同完成用户拿取/放回商品的行为动态检测，以及多商品重叠场景下的识别任务；用户结算时，系统再将视觉识别结果与传感器数据进行交叉验证，确保结算的精准性。从实际落地效果来看，该系统能同时处理多个用户的购物行为，精准识别用户拿取的商品，支撑了“即拿即走”的无结算体验；这一方案的技术选型逻辑，也成为后续同类无人零售场景的标准参考方案。

7-Eleven（日本）：其技术方案的核心逻辑，是“轻量化边缘计算设备 + 高精度模型”的组合，聚焦于实现低成本、高实时性的货架商品检测 —— 这一方案的技术选型逻辑，更适配传统零售业态的升级需求。算法层面，选择了轻量化的 YOLOv5s 模型作为骨干检测框架，对货架上的商品进行实时检测，生成商品陈列状态的实时数据；为了适配边缘设备的算力约束，方案采用了模型蒸馏、边缘计算等多重优化手段，将模型的参数量压缩了近 70%，同时保留了核心的特征提取能力。从实际落地效果来看，该系统的商品定位精度达到 92.7%，单排货架的扫描耗时控制在 200 毫秒以内；支撑了后台系统实时生成货架的陈列状态、商品库存数量、缺货位置等核心运营数据；这一技术方案的落地，将单店的人工盘点效率提升了约 80%，同时将人工巡检的频率压缩了 90%，大幅降低了门店的运营成本。

国内某头部电商企业：其技术方案的核心逻辑，是“多模型融合 + 轻量化部署”，支撑全渠道场景下的商品数字化识别任务。算法层面，综合了两阶段检测

器 Faster R-CNN 与单阶段检测器 YOLOv8 的优势，设计了并行的多模型融合架构——由 YOLOv8 完成初步筛选，快速定位潜在目标区域；再由 Faster R-CNN 完成精细分析，通过置信度加权融合机制整合两者的检测结果。在部署环节，采用了轻量化模型设计、GPU 推理加速、边缘计算等优化技术，将不同的检测方案，适配到不同的边缘算力设备上。从实际落地效果来看，融合后的方案平均准确率达到 82.5%，单张图片处理时间为 0.3 秒，在保证实时性的前提下，大幅提升了识别精度；这一技术方案的落地，将商家的商品上架标注效率提升了 40%，同时将库存盘点的人工成本降低了约 70%。

从上述案例可以看出，商品识别技术的落地选型逻辑，完全由实际业务场景的技术约束决定——没有最先进的技术方案，只有最适配的选型：对实时性要求高的场景，优先选择单阶段检测器；对精度要求高的场景，优先选择两阶段检测器；当场景技术约束多元、复杂时，多模型融合是当前最优的平衡方案。

5.2 虚拟试穿行业应用案例

与商品识别的间接价值不同，虚拟试穿的技术落地效果，能直接体现在用户体验和业务核心指标上——头部平台的公开数据，也直观验证了这一技术的商业价值与落地可行性。本文选取了四个行业内具有代表性的落地案例，从技术选型、场景适配及实际业务效果的维度，进行详细的对比分析：

京东 AI 试穿 Oxygen Tryon（中国）：技术选型逻辑完全贴合电商场景的用户使用习惯——从第一代基于 Diffusion+U-Net 架构的方案，迭代到第四代的 Mask Free+DiT 架构，全程采用自研技术框架。其中，第二代方案将扩散模型的主体架构从 U-Net 切换为 DiT，同时用 ViT 与 VAE 编码器完成特征提取，实现了服装细节、人体姿态的精准融合；第四代方案引入视觉语言模型（VLM）进行多模态特征表征，进一步简化了技术流程，提升了多件搭配场景的适配能力。在部署环节，通过模型蒸馏、轻量化裁剪等技术优化，将高分辨率场景下的推理时间控制在 7 秒以内。该技术方案落地后，成为京东服饰品类的核心流量入口，有效降低了退货率，提升了用户购买转化率。

阿里 Lookie 与淘宝 AI 试穿（中国）：采用“双产品协同”的技术布局逻辑，

通过两款互补的产品，分别覆盖不同的用户场景需求，平衡试穿效果与用户使用成本。其中，淘宝 AI 试穿聚焦单件商品的试穿核心场景，采用 Diffusion+DiT 架构的方案，直接嵌入到商品搜索商卡的核心流量入口，用户无需额外跳转，就能快速获取商品的上身效果，以极简的操作链路覆盖大多数用户的核心需求；Lookie 则聚焦多件搭配、具有娱乐属性的场景，采用 Diffusion+U-Net 架构的方案，用户只需上传几张照片就能生成专属数字分身，基于分身完成多件搭配试穿，覆盖用户组合穿搭、场景化搭配的延伸需求。这一方案的落地，将阿里服饰类商品的用户试穿后转化率提升了显著幅度，成为平台降低退货率、提升用户停留时长的核心技术支撑。

Zara（西班牙）：采用“APP 端 + 线下门店端”的多端协同技术落地逻辑，将虚拟试穿作为连接线上、线下消费场景的核心技术枢纽，覆盖不同渠道用户的体验需求——线上通过 APP 的 AR 试穿功能，采用 Diffusion+U-Net 架构的方案，用户只需上传全身照片，就能在短时间内获取试穿效果图；线下门店则提供基于深度摄像头和 AR 技术的智能试衣镜，用户在镜子前做出动作，试衣镜内的虚拟模特就会同步展示服饰的上身效果。这一技术方案落地后，官网数据显示，服饰类商品的在线销售退货率从 35% 降至 22%，用户的试穿后转化率提升了近 30%。

Amazon Fashion（美国）：技术布局的核心逻辑，是“虚拟试穿+个性化定制”的全链路打通——技术选型以 Diffusion+DiT 架构为基础，结合自身在 3D 人体扫描、深度学习上的技术积累，开发了适配全品类的专属试穿方案。用户使用时，先通过鞋类的 AR 试穿功能，扫描脚部的 13 个关键数据点，获取精准的脚步尺寸数据；这些尺寸数据会同步到服饰类商品的试穿环节，结合用户的身材维度数据，生成更贴合真实身形的试穿效果。这一方案落地后，Amazon Fashion 的鞋类商品退货率直接下降 25%，成为驱动业务增长的重要技术支撑。

可以看到，头部电商平台的虚拟试穿技术选型逻辑，完全以“用户体验与业务效果的平衡优先级”为核心：对用户使用门槛要求高的场景，优先选择轻量化、易落地的方案；对试穿效果要求高的核心场景，优先选择生成质量更强、细节还原度更高的扩散模型+DiT 架构方案；当场景需要覆盖多维度、复杂的用户需求时，多方案组合协同是当前最优的平衡选择。

6 结论

随着电子商务的体验化、智能化转型加速，深度学习驱动的计算机视觉技术，已经从行业的“可选配置”升级为“必备基础设施”——其技术价值的核心逻辑，是通过视觉信息的数字化，打通“人-货-场”的全链路数字化闭环：在商品识别端，替代人工完成高频、重复的商品数字化管理工作，大幅提升运营效率；在虚拟试穿端，将平面的商品图像，转化为具备真实感的个性化场景化体验，补齐线上购物的体验短板。

从技术演进的维度来看，两类技术的路线选择，已经完成了从“传统人工特征提取”到“深度学习端到端特征提取”的完整迭代；技术选型的逻辑，已经从“单纯追求精度或速度”升级为“基于场景化技术约束的多维度平衡选型”。经过头部电商平台的大规模落地验证，以扩散模型、Transformer 架构、多模态融合为核心的技术方案，是当前匹配电商场景技术约束的最优选择，也是后续技术升级的基础底座。

从行业的技术进展维度来看，目前技术的成熟度，已经足以支撑企业级规模化部署，但仍面临四个需要突破的共性瓶颈：一是高轻量化与高精度性的平衡瓶颈；二是训练数据的瓶颈约束；三是复杂场景下的鲁棒性瓶颈；四是多模态融合的细节级对齐瓶颈。这些瓶颈的存在，共同限制了技术的效果上限，也成为行业后续技术升级的重点方向。

可以预见，随着轻量化技术、多模态大模型技术、生成式 AI 技术的持续迭代升级，以及行业内场景化数据的逐步积累，计算机视觉技术将在电子商务行业得到更深层次的普及应用，从“支撑业务运营的技术工具”演进为“驱动业务增长的核心基础设施”——不仅会持续优化消费者的购物体验，更会进一步重构电商行业的供应链链路，从技术层面支撑电商行业的效率提升与业务增长。

参考文献

- [1] 彭翔月。基于 Faster R-CNN 与 YOLOv8 多模型融合的电子商务图像分类与识别[J].电子商务评论, 2025,14 (3):306-313.
- [2] 高继航。迈向电商大模型时代: 从虚拟试穿到电商 AIGC [EB/OL]. 51CTO,2026.
- [3] 李岩。京东 Oxygen Tryon: AI 驱动的虚拟试穿技术实践 [EB/OL]. 京东云官方技术博客, 2025.
- [4] 张成.LSR-YOLO: 轻量化零售商品检测模型 [J]. PMC,2026.
- [5] 方佳乐. 计算机视觉技术在人工智能领域的应用研究 [J]. 中国高新技术,2021,(20):40-41.DOI:10.13535/j.cnki.10-1507/n.2021.20.06.
- [6] 李翔. 浅谈计算机视觉及其在电子商务中的应用 [J]. 商场现代化,2020,(4):35-36.DOI:10.14013/j.cnki.scxdh.2020.04.014.
- [7] 刘冰,胡丽娜.数字技术赋能电子商务与新质生产力融合发展研究[J].中国电子商情,2026,32(11):7-9.
- [8] 孙华东,电子商务平台商品图像分类关键技术研究.黑龙江省,哈尔滨商业大学,2022-07-13.
- [9] 杨蕊.基于内容的电子商务图像分类研究[D].西北大学,2015.
- [10]宋贤霞.基于内容的图像检索在电子商务中的应用[J].电脑知识与技术,2012,8(15):3744-3745+3754.
- [11]张春艳.电商网站中面料商品的情感自动分类研究[D].北京服装学院,2017.
- [12]樊慧琼.基于深度学习的电子商务产品数据分析研究[D].江西财经大学,2025.DOI:10.27175/d.cnki.gjxcu.2025.000036.
- [13]杜雅卓.多媒体技术在电子商务网站中的应用研究[J].信息记录材料,2022,23(11):117-119.DOI:10.16009/j.cnki.cn13-1295/tq.2022.11.053.
- [14]魏稼熒.电子商务中视觉搜索的图像表征方法研究[D].浙江大学,2022.DOI:10.27461/d.cnki.gzjdx.2022.000066.
- [15]夏洪起.电子商务中图像视觉效应对消费者参与影响研究[J].中国市场,2021,(27):71-72.DOI:10.13939/j.cnki.zgsc.2021.27.071.