MULTI-MODAL SENSING FOR REAL-TIME OPERATIONS MONITORING IN PHYSICAL RETAIL: A FRAMEWORK COMBINING COMPUTER VISION, IOT TELEMETRY, AND TRANSACTIONAL SIGNALS

Kemal BUYUKTAS
Flo Technology, Technology Department, Istanbul, Türkiye
ORCID: 0009-0007-7884-2872

ABSTRACT

Physical retail stores continue to generate a substantial share of total revenue in many categories, but the operational visibility available to store managers still lags far behind what digital channels provide. Workforce activity, opening and closing routines, customer density, checkout flow, fitting-room dynamics, and after-hours store activity are typically inferred from manual observation or sales proxies rather than measured continuously. As a consequence, structural inefficiencies often remain invisible until they show up as lost sales or compliance incidents. This study reports on a multi-modal sensing framework that narrows this gap by re-purposing in-store infrastructure already in place for unrelated reasons: security cameras, point-of-sale terminals, door and turnstile counters, energy and lighting sensors, fitting-room occupancy signals, and the corporate time-attendance system. Signals from these sources feed a shared operational layer connected to a real-time alerting and feedback mechanism. The framework was deployed in twelve pilot stores of a large omnichannel apparel and footwear retailer over a six-month window, under a privacy-preserving, edge-inference architecture. A directional, observational analysis indicates that idle staff time during peak hours fell by around 20%, coverage of previously underserved zones grew by close to 30%, unrecorded early closures dropped sharply (about 81%), average checkout wait times shortened by roughly a quarter, fitting-room abandonment decreased by approximately 20%, and staff time-attendance discrepancies fell by close to 60%. Two stores diverged from this pattern, illustrating that camera placement and sensor coverage materially affect what such a framework can deliver.

Keywords: Multi-Modal Sensing, Retail Operations, Computer Vision

1. INTRODUCTION

Despite the continued expansion of e-commerce, physical retail stores remain the channel through which most customers in apparel and footwear actually complete purchases. In these categories, product fit, material perception, and immediate availability are still decisive, and the strategic question for retailers has shifted from whether to invest in physical stores to how to operate them at a level of granular efficiency closer to that of digital channels (Verhoef et al., 2021; Grewal et al., 2022).

The main obstacle to that ambition is uneven visibility between the two worlds. Digital channels generate event-level data on every interaction; physical stores still depend mostly on transaction logs, periodic audits, and supervisor judgment. What happens between transactions — where customers cluster, where staff are positioned relative to demand, how long someone waits at checkout, whether a fitting room is ever attended, or whether the store actually opened on time — is rarely measured directly (Pantano et al., 2020; Caro et al., 2020). When such signals are absent, inefficiencies tend to persist until they finally show up as lost sales, complaints, or audit findings.

Most stores already operate a heterogeneous set of digital systems, but for purposes other than continuous operational analytics. Cameras are installed for loss prevention. Point-of-sale terminals are operated for transactions. Door and turnstile counters produce footfall reports. Energy and lighting systems are managed for cost. Fitting-room occupancy mechanisms exist for service routing. Time-attendance systems are operated for payroll. Each of these systems on

its own shows only a narrow slice of reality. When their signals are integrated into a shared analytical layer and coupled with intervention mechanisms, the existing infrastructure can collectively act as a continuous operational signal layer without additional hardware investment (Sestino et al., 2020; Bertsimas & Kallus, 2020).

The retail-analytics literature has so far engaged with these modalities mostly one at a time. Computer vision research at the store level is typically constrained to entrance footfall counting or theft detection. Point-of-sale studies treat transactional data as a marketing or promotional signal rather than as an operational one. IoT-in-retail work tends to focus on a single device class. The systematic combination of these modalities into an intervention-oriented operational layer, together with the operational, ethical, and privacy questions that arise once they coexist, has received much less attention.

The work reported here is intended to close that gap. The paper proposes a multi-modal sensing framework for real-time operations monitoring in physical retail and reports a directional, observational evaluation in twelve pilot stores. The framework combines vision-based workforce activity tracking, cross-modal opening and closing discipline detection, customer density and queue analytics, fitting-room utilization monitoring, after-hours activity detection, and staff time-attendance reconciliation. It is designed to run on existing infrastructure, to keep raw signals on-premises, and to produce metrics that store managers can act on. Section 2 reviews the relevant literature; Section 3 describes the methodology and architecture; Section 4 reports the directional results; Section 5 discusses interpretation, managerial implications, and limitations; Section 6 concludes.

2. LITERATURE REVIEW

Verhoef et al. (2021) argue that digital transformation in retail has stopped being optional and that physical stores must close the visibility gap relative to digital channels to remain competitive. Grewal et al. (2020) take a similar position, pointing out that the future of retailing depends on continuous instrumentation of customer interactions across both physical and digital touchpoints. In a later contribution, Grewal et al. (2022) name camera-based analytics, artificial intelligence, and edge computing as the technologies with the highest near-term impact on store operations. Caro et al. (2020) shift the emphasis toward operations, arguing that the future of retail operations will be defined by closer integration between operational decisions and continuously updated in-store data, not by methodology choice alone. Hoyer et al. (2020) reinforce this from the customer side, observing that customer experience in the digital age is shaped by continuous data integration and consistent treatment of customer signals across channels; this is precisely what motivates a continuous in-store sensing layer. Pantano et al. (2020) document a more recent pressure: COVID-19 sharply accelerated retailer demand for granular in-store occupancy and density signals, making this kind of instrumentation a near-term operational concern rather than a longer-term research direction.

The technical foundation for in-store visual instrumentation has matured rapidly. Carion et al. (2020) show that Transformer-based detection architectures provide reliable person detection at real-time frame rates without the hand-engineered components of earlier pipelines. Wang et al. (2023) report that contemporary YOLO-family detectors achieve state-of-the-art performance with substantially lower inference latency, which makes them appropriate for store-level edge devices. Zhang et al. (2022) find that modern multi-object trackers such as ByteTrack maintain identity assignment robustly under occlusion, a property of obvious relevance in dense retail environments. Aharon et al. (2022) take this further, demonstrating that hybrid motion-and-appearance association strategies reduce identity switches still further, and Ye et al. (2022) point out that deep-learning-based person re-identification supports cross-camera linkage at accuracy levels sufficient for store-scale operational analytics. Dogan and

Akyuz (2020) take a more applied angle, showing that indoor customer behavior analysis based on in-store cameras can yield actionable layout and staffing insights, provided it is treated as an operational system rather than a one-off study.

Beyond cameras, the maturation of IoT, edge computing, and prescriptive analytics has made multi-modal in-store sensing feasible. Sestino et al. (2020) argue that the Internet of Things and big data together act as enablers of business digitalization strategies, including the integration of heterogeneous operational signals into a shared analytical layer. Cao et al. (2020) report that edge computing reduces bandwidth requirements, latency, and exposure of raw data, which makes it a natural substrate for visual and sensor analytics in privacy-sensitive contexts. Bertsimas and Kallus (2020) reframe the analytics question itself: they argue that the productive move is from predictive to prescriptive analytics, in which analytical outputs are coupled directly with operational decisions. That reframing motivates building closed-loop intervention systems rather than passive dashboards.

The ethical and regulatory side has, finally, started to receive sustained attention. Quinn (2021) shows that data-protection-aligned processing of identifiable visual and sensor data requires explicit purpose limitation, data minimization, and demonstrable proportionality, particularly in employee-monitoring contexts. On the methodological side, Vom Brocke et al. (2020) present design science research as a principled framework for constructing operational artifacts under combined performance, organizational, and regulatory constraints.

Taken together, the literature still leaves a few related gaps. In-store sensing modalities (computer vision, point-of-sale analytics, IoT telemetry, fitting-room signals, energy management, and staff time-attendance systems) are typically studied in isolation rather than as components of a single, integrated operational framework. Empirical results from active retail environments are rare; most published evaluations come from controlled or laboratory settings. And privacy-preserving architectural choices are still mostly treated as a compliance layer placed on top of an existing pipeline, rather than as a foundational design property. The work reported here is structured to address these gaps.

3. METHODOLOGY

3.1 Approach and Pilot Context

The framework was developed and evaluated using a design science research approach (Vom Brocke et al., 2020): the artifact itself is the object of study, and it was built iteratively, deployed progressively, and measured against a small set of pre-defined operational outcomes, rather than tested against statistical hypotheses in a controlled setting.

The pilot involved twelve physical retail stores operated by a large omnichannel apparel and footwear retailer in Türkiye. Stores varied in floor area, layout, and average daily traffic, which exposed the pipeline to heterogeneous operational conditions. The observation window covered six months, including both low-season weekdays and weekend campaign peaks. Two stores were partially excluded from individual analyses: in one case a four-week POS upgrade prevented reliable queue-time inference, and in the other camera coverage had been changed mid-pilot for unrelated security reasons. These exclusions are flagged where relevant in Section 4.

The deployment reused the digital systems that were already in place: security cameras (originally installed for loss prevention), point-of-sale terminals, electronic door and turnstile counters, energy and lighting telemetry, fitting-room occupancy sensors and call buttons, and the corporate time-attendance system. Each store was additionally equipped with an on-premises edge inference device to run the computer vision pipeline locally and to fuse the resulting signals with structured data from the other systems.

3.2 Architecture

The framework is organized as a unified operational signal layer. Heterogeneous sensors and back-office systems feed a common analytical core, and the core in turn drives a real-time alerting and feedback mechanism. Seven modules contribute to the signal layer: vision-based workforce activity tracking, cross-modal opening and closing discipline detection, customer density and queue analytics, fitting-room utilization monitoring, after-hours activity detection, staff time-attendance reconciliation, and a privacy-preserving aggregation layer that mediates all data flows. The unified store operations data layer aggregates module outputs and exposes them through dashboards and notification channels used by store managers and regional supervisors. Video and raw sensor data are processed locally; only aggregated, non-identifying metrics are transmitted to the central platform.

3.3 Workforce Activity Tracking (Computer Vision)

Person detection runs on a YOLO-family detector tuned for in-store viewpoints, followed by a DeepSORT-style tracker that maintains consistent track identifiers across frames. A lightweight role classifier distinguishes staff from customers using uniform appearance features rather than biometric identification; misclassified frames are absorbed through temporal smoothing. During the first two weeks of deployment, the role classifier confused customers wearing dark monochrome clothing with staff, particularly under low-light evening conditions; this was addressed by retraining the classifier on additional in-store imagery. Per-shift, per-zone activity indicators are computed for staff, including total active time, idle-time intervals, seated-time intervals, and coverage of designated service zones. The system tracks roles and locations, not identities; no individual employee or customer is identified, persisted, or re-identifiable across visits.

3.4 Opening and Closing Discipline (Cross-Modal Validation)

Opening and closing discipline detection compares observed staff and customer presence against scheduled store hours, drawing on three independent signal sources: vision-based presence detection, door and turnstile telemetry, and point-of-sale terminal activity. Late openings, premature closings, and intervals during which the store appears unattended within scheduled hours are flagged only when at least two of the three sources agree, which sharply reduces the false-positive rate compared with any single source. Detections are surfaced as ranked, attributable events tied to a specific time window and store zone, rather than as global red or green indicators.

3.5 Customer Density and Queue Analytics

Customer density heatmaps are built by projecting detected customer positions onto a calibrated floor plan and aggregating presence over rolling time windows. The output identifies high-traffic zones, underserved areas, and emerging clusters around specific product categories or service points. Density is computed at an aggregate level only; individual trajectories are not retained. Checkout queue length and wait time are estimated from a hybrid signal that combines vision-based queue length detection with point-of-sale inter-transaction intervals, which gives a more stable estimate than either modality on its own.

3.6 Fitting-Room Utilization

Fitting-room utilization is monitored through occupancy sensors and call-button telemetry, supplemented by computer vision-based entry and exit signals at the fitting-room corridor. The module computes utilization rate, average occupancy time, abandonment events (customers who enter the corridor but do not enter a fitting room), and call-button response time. In apparel and footwear, fitting rooms are conversion-critical touchpoints, so this module is treated as a core operational signal rather than as a peripheral metric.

3.7 After-Hours Activity Detection

After-hours activity detection identifies unauthorized presence or operational activity outside scheduled store hours by combining energy consumption, lighting state, door telemetry, and computer vision. A flag is raised when at least two independent modalities indicate activity inconsistent with the closed-store baseline. This module addresses a class of incidents that is essentially invisible to traditional security monitoring, which is built around intrusion alarms rather than operational deviation.

3.8 Staff Time-Attendance Reconciliation

Staff time-attendance reconciliation cross-checks the corporate time-attendance system against on-premises presence signals derived from computer vision and access-control logs. Discrepancies, for example recorded clock-ins without corresponding presence or observed presence without corresponding clock-ins, are surfaced as ranked events for managerial review. The module operates on aggregate, role-level signals, not on identity-level tracking, and the corporate time-attendance system remains the system of record.

3.9 Privacy-Preserving Design

Privacy was treated as a design constraint from the beginning, not as a compliance layer added later. All video processing runs on the edge device; raw frames are not persisted to disk and are not transmitted outside the store. Sensor and transactional signals are aggregated at the store level before transmission. Only aggregated, non-identifying metrics (counts, durations, densities, queue lengths, and event flags) leave the device. No biometric features are stored, no facial recognition is performed, and no individual is re-identifiable across sessions. The deployment was reviewed against national data protection regulations, and employees were informed about both the operational purposes and the non-individual nature of the analytics.

3.10 Alerting and Feedback

The alerting and feedback layer turns module outputs into specific managerial actions. Typical alerts include a notification to a store manager when a high-density zone has been undercovered by staff for more than a defined duration, a daily compliance summary on opening and closing, a real-time alert when a checkout queue exceeds a configurable wait threshold, an immediate flag when after-hours activity is detected, and an automated suggestion for staff redeployment based on the joint distribution of customer density and current staff coverage. Outcomes from those interventions are written back into the analytics layer, which means detection thresholds can be refined progressively over the life of the deployment.

3.11 Deployment and Measurement

Deployment followed a three-phase rollout. The first phase ran the perception and sensor-fusion pipelines in shadow mode to validate signal quality without surfacing alerts. The second phase activated dashboards and daily summaries so that managers could compare the multi-modal signal with their own operational expectations. The third phase enabled real-time alerting, staff redeployment suggestions, and integration with regional compliance reviews. Nine operational outcome metrics were defined in advance, covering workforce, customer experience, compliance, and energy efficiency, and pre- and post-deployment values were compared across the participating stores.

4. RESULTS

The results below come from a directional, observational analysis of pilot deployment data collected across twelve participating stores during a six-month window that covered both weekday and weekend traffic patterns and both low-season and promotional peaks. The deployment was non-randomized and embedded in active retail operations, so the analysis supports directional inference rather than causally identified effect estimation; this is an

acknowledged limitation of the study (Section 5). Table 1 summarizes the observed pre- and post-deployment values for the nine operational outcome metrics.

Table 1. Pre- and post-deployment values for the operational outcome metrics

Outcome metric	Pre-deployment baseline	Post-deployment value	Directional change
Idle staff time during peak operating hours	28.4% of staff-hours	22.6% of staff-hours	20.4% reduction
Coverage of underserved store zones	41.0% of expected coverage	53.2% of expected coverage	29.8% increase
Unrecorded early closures	3.7 events per store per month	0.7 events per store per month	81.1% reduction
Average checkout queue wait time (peak hours)	138 seconds	104 seconds	24.6% reduction
Fitting-room abandonment rate	16.3%	13.1%	19.6% reduction
After-hours unauthorized activity events captured	Not observable	14 events	New visibility
Energy consumption during unstaffed intervals (indexed)	100.0	84.7	15.3% reduction
Staff time-attendance discrepancies	9.1% of shifts	3.6% of shifts	60.4% reduction
Conversion rate in high-density zones (peak hours)	18.2%	19.7%	8.2% increase

Note. *Figures come from a directional, observational analysis of pilot deployment data across twelve participating stores over a six-month window. They should not be interpreted as causally identified effect estimates.*

Two stores diverged from the broader pattern. In one case, suboptimal camera placement reduced detection reliability for the workforce activity module; in the other, a mid-pilot POS upgrade interrupted the queue-time signal for four weeks. Both stores remain in the aggregate figures except where explicitly noted.

4.1 Idle Staff Time

Across the participating stores, idle staff time during peak operating hours moved from 28.4% of total staff-hours before deployment to 22.6% after deployment, a directional reduction of about 20%. The shift was most pronounced in the larger stores, where supervisor visibility had previously been limited and idle pockets had tended to persist undetected. In interviews, store managers consistently linked the change to the redeployment suggestions generated by the alerting layer rather than to a sense of being monitored, which is consistent with the design intent of treating workforce signals as decision support rather than performance scoring.

4.2 Coverage of Underserved Zones

The framework's largest measurable impact emerged in coverage of underserved zones, those chronically understaffed relative to their average customer density. Average coverage rose from 41.0% to 53.2% of expected coverage, a directional increase of close to 30%. Zones that had previously fallen below dashboard visibility entered the redeployment shortlist routinely after Phase 3.

4.3 Unrecorded Early Closures

Once cross-modal validation went live, unrecorded early closures (intervals within scheduled store hours during which the store appeared unattended without an authorized exception) fell from an average of 3.7 events per store per month to 0.7, a directional reduction of around 81%. The cross-modal opening and closing module proved particularly useful here, because requiring agreement among multiple sources substantially reduced the false-positive rate that had previously eroded managerial trust in single-source compliance signals.

4.4 Checkout Queue Wait Time

Average checkout queue wait time during peak hours moved from 138 seconds to 104 seconds, a directional reduction of about a quarter. The hybrid queue signal, which combines vision-based queue length detection with point-of-sale inter-transaction intervals, generated earlier and more reliable cashier-redeployment alerts than either modality on its own. One store was excluded from this metric because of the POS upgrade noted above.

4.5 Fitting-Room Abandonment

Fitting-room abandonment, defined as the share of customers who entered the corridor but did not complete a fitting-room session, fell from 16.3% to 13.1%, roughly a 20% reduction in directional terms. Counter to initial expectations, the change was actually larger in smaller stores than in larger ones. In informal interviews, managers attributed this to the fact that smaller stores had previously relied on incidental staff presence near the corridor, which is exactly the kind of slack the new call-button alerts removed.

4.6 After-Hours Activity and Energy Use

Fourteen after-hours unauthorized activity events were captured across the pilot stores during the observation window. These had been structurally invisible to traditional security monitoring, which is built around intrusion alarms rather than operational deviation. In parallel, energy consumption during unstaffed intervals fell by about 15% on an indexed basis, as managers responded to cross-modal evidence that lighting and HVAC had been left active during low-activity windows. The energy improvement is reported here as a side-effect of the framework rather than as a primary target.

4.7 Staff Time-Attendance Discrepancies

Staff time-attendance discrepancies, defined as recorded clock-ins without corresponding presence or observed presence without corresponding clock-ins, fell from 9.1% of shifts to 3.6%, a directional reduction of about 60%. The improvement came primarily from visibility and managerial follow-up rather than from automated escalation; the corporate time-attendance system was kept as the system of record throughout.

4.8 Conversion in High-Density Zones

Conversion rate in zones identified as high-density during peak hours rose from 18.2% to 19.7%, a directional lift of roughly 8%. The lift was consistent in direction across stores, which supports the interpretation that it is associated with the framework's interventions; we do not, however, make a causal claim, given the observational nature of the analysis.

5. DISCUSSION

5.1 Interpretation

The directional results lend support to a few interrelated propositions. The first is that in-store operations efficiency is driven less by total headcount than by the joint distribution of workforce activity, customer density, compliance, and infrastructure state. The reductions in idle staff time and the increases in zone coverage were achieved without changes in headcount. Caro et al. (2020) make a closely related argument: that the future of retail operations is defined by tighter integration between operational decisions and continuously updated data, not by static methodology choice. Hoyer et al. (2020) frame this from the customer side, arguing that continuous data integration is a precondition for differentiated customer experience; the framework's effect on checkout queues and fitting-room abandonment is consistent with that argument.

The second proposition is that multi-modal sensing is most effective when it is treated as a single layer with intervention attached, rather than as a portfolio of single-modality dashboards. That is essentially the move from predictive to prescriptive analytics described by Bertsimas and Kallus (2020), translated into a multi-modal retail setting. The opening and closing discipline result is the clearest illustration: a single-source compliance signal that had previously eroded managerial trust because of false positives became actionable once at least two of three independent sources had to agree.

A third proposition concerns privacy. The framework was built so that no biometric features are stored, no individual is re-identifiable across sessions, and aggregated metrics are the only data that ever leave the store. The fact that operational outcomes still moved across nine metrics suggests that meaningful retail analytics does not require identifying anyone, which supports the position taken by Quinn (2021) that data-protection-aligned processing can coexist with substantive operational insight.

5.2 Managerial Implications

A few implications follow for retail operations leaders. The first is that existing in-store infrastructure (cameras, POS terminals, door and turnstile counters, energy telemetry, fitting-room sensors, time-attendance systems) can be re-purposed into a continuous operational signal layer with no additional hardware investment, provided the integration is designed at the same time as the perception layer. The second is that cross-modal validation produces signals managers are willing to act on, which matters more for compliance-sensitive events than for descriptive analytics, because trust in the signal directly determines the speed and consistency of managerial follow-up. The third is that density-aware staff redeployment and queue-aware cashier redeployment are higher-leverage moves than incremental headcount adjustments, particularly during peak hours and in larger formats. A fourth, more secondary observation across stores is that compliance and energy outcomes tend to improve through visibility alone rather than through escalation, which has direct implications for the design of regional performance management routines and for the role of district managers.

5.3 Limitations

Several limitations should be acknowledged. The pilot was conducted within a single retailer's store network, and outcomes may differ for retailers operating with markedly different formats, customer profiles, or supervisory structures. The deployment was non-randomized and embedded in active operations, so the analysis is directional and observational rather than causally identified; some of the observed change may reflect concurrent operational adjustments rather than the framework alone. Two stores were partially excluded from individual metrics due to camera placement and POS upgrade issues, and these exclusions are individually noted in Section 4. Cross-modal fusion also introduces a dependency on the

reliability of multiple upstream systems; failure or drift in any one source can degrade the unified signal, which is something to monitor over longer deployment horizons. Finally, longitudinal effects on employee experience, customer satisfaction, and store-level retention were not directly measured in this study and remain priorities for future work.

6. CONCLUSION

This paper has described a multi-modal sensing framework for real-time operations monitoring in physical retail, combining computer vision, point-of-sale analytics, door and turnstile telemetry, energy and lighting sensors, fitting-room occupancy signals, and staff time-attendance records into a unified analytical layer connected to a real-time alerting and feedback mechanism. The framework was designed to run on existing in-store infrastructure, to keep raw signals on-premises, and to produce metrics that managers can act on without violating the privacy expectations of customers or staff.

A directional, observational analysis of pilot deployment data from twelve stores over six months indicates that idle staff time during peak hours fell by around 20%, coverage of previously underserved zones grew by close to 30%, unrecorded early closures dropped by roughly 81%, average checkout wait times fell by about a quarter, fitting-room abandonment decreased by around 20%, energy use during unstaffed intervals fell by approximately 15%, time-attendance discrepancies fell by close to 60%, and conversion in high-density zones rose by approximately 8%. Fourteen after-hours unauthorized activity events that had previously been invisible to traditional monitoring were captured during the observation window.

The central claim of the paper is therefore a modest but specific one: the operational visibility gap between physical and digital retail can be narrowed without new hardware investment, provided that heterogeneous in-store signals are integrated, coupled with intervention mechanisms, and governed by a privacy-preserving design from the start. Future work might usefully extend the framework to multi-retailer settings, examine cross-format variation in customer-staff dynamics, incorporate causally identified evaluation designs to complement the directional evidence reported here, and add a longitudinal layer that tracks employee experience and customer satisfaction over time alongside the operational metrics.

REFERENCES

Aharon, N., Orfaig, R., & Bobrovsky, B. Z. (2022). BoT-SORT: Robust associations multi-pedestrian tracking. arXiv preprint, arXiv:2206.14651.

Bertsimas, D., & Kallus, N. (2020). From predictive to prescriptive analytics. Management Science, 66(3), 1025–1044.

Cao, K., Liu, Y., Meng, G., & Sun, Q. (2020). An overview on edge computing research. IEEE Access, 8, 85714–85728.

Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-end object detection with Transformers. Proceedings of the European Conference on Computer Vision (ECCV), 213–229.

Caro, F., Kök, A. G., & Martínez-de-Albéniz, V. (2020). The future of retail operations. Manufacturing & Service Operations Management, 22(1), 47–58.

Dogan, O., & Akyuz, A. O. (2020). Indoor customer behavior analysis with computer vision techniques. International Journal of Retail & Distribution Management, 48(11), 1149–1165.

Grewal, D., Roggeveen, A. L., & Nordfält, J. (2020). The future of retailing. Journal of Retailing, 96(1), 69–85.

Grewal, D., Motyka, S., & Levy, M. (2022). The future of retailing: Emerging trends and technologies. Journal of Retailing, 98(1), 1–10.

Hoyer, W. D., Kroschke, M., Schmitt, B., Kraume, K., & Shankar, V. (2020). Transforming the customer experience through new technologies. Journal of Interactive Marketing, 51, 57–71.

Pantano, E., Pizzi, G., Scarpi, D., & Dennis, C. (2020). Competing during a pandemic? Retailers' ups and downs during the COVID-19 outbreak. Journal of Business Research, 116, 209–213.
Quinn, P. (2021). Research under the GDPR – A level playing field for public and private sector research? Life Sciences, Society and Policy, 17(1), 4.
Sestino, A., Prete, M. I., Piper, L., & Guido, G. (2020). Internet of Things and big data as enablers for business digitalization strategies. Technovation, 98, 102173.
Verhoef, P. C., Kannan, P. K., & Inman, J. J. (2021). Digital transformation in retail. Journal of Retailing, 97(1), 1–17.
Vom Brocke, J., Hevner, A., & Maedche, A. (2020). Introduction to design science research. In J. Vom Brocke, A. Hevner, & A. Maedche (Eds.), Design science research cases (pp. 1–13). Springer.
Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 7464–7475.
Ye, M., Shen, J., Lin, G., Xiang, T., Shao, L., & Hoi, S. C. H. (2022). Deep learning for person re-identification: A survey and outlook. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(6), 2872–2893.
Zhang, Y., Sun, P., Jiang, Y., Yu, Z., Yuan, Z., Luo, P., Liu, W., & Wang, X. (2022). ByteTrack: Multi-object tracking by associating every detection box. Proceedings of the European Conference on Computer Vision (ECCV), 1–21.