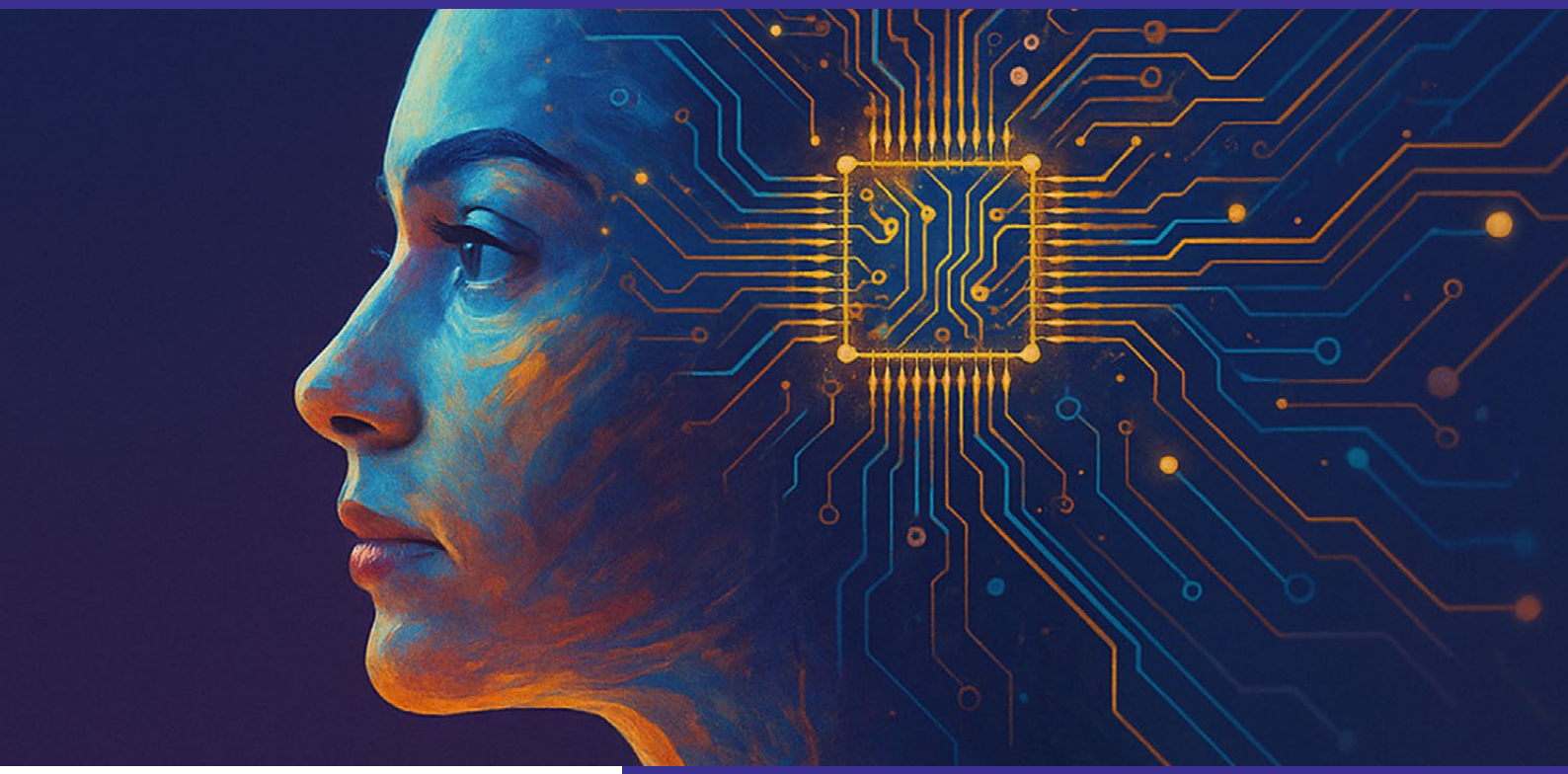
ESCP
BUSINESS SCHOOL

IT ALL STARTS HERE

BERLIN | LONDON | MADRID | PARIS | TURIN | WARSAW



LIGHTS - Technology

Computer vision empowering brand strategy: from vision to action

ESCP Impact Paper No.2025-34-EN

Yaqiu LI, Hsin-Hsuan Meg LEE & Lorena BLASCO-ARCAS
ESCP Business School

ESCP RESEARCH INSTITUTE OF MANAGEMENT (ERIM)

ESCP Impact Paper No.2025-34-EN

Computer vision empowering brand strategy: from vision to action

Yaqiu Li*
Hsin-Hsuan Meg Lee**
Lorena Blasco-Arcas**

ESCP Business School

Abstract

In today's visually driven digital landscape, computer vision has become an invaluable asset for brand managers seeking deeper consumer insights and more effective brand strategies. By integrating low-level feature extraction (e.g., color and texture analysis) with higher-level tasks (e.g., object detection, facial expression recognition, and image generation), brands can systematically assess and enhance critical visual elements such as logos, product packaging, and user-generated content. Drawing upon consumer-based brand equity (CBBE) principles, this paper illustrates how computer vision tools can strengthen brand identity (e.g., detecting copycats), deepen brand meaning (e.g., uncovering hidden associations from social images), capture real-time consumer response (e.g., measuring emotion), and nurture brand relationships (e.g., understanding engagement). We also highlight emergent possibilities for predictive analytics, personalization, and rebranding by fusing computer vision with multimodal data, text, audio, and videos. Effectively leveraging computer vision can significantly boost efficiency and unlock new opportunities for managers and strategic innovation in today's highly competitive, visually driven marketplace.

Keywords: computer vision, brand strategy, image analysis, artificial intelligence

*Doctoral researcher, ESCP Business School
**Professor, ESCP Business School

Computer vision empowering brand strategy: from vision to action

Introduction

In the digital economy era, brand competition is rapidly shifting from being "text-centric" to "visually driven." Whether in a TikTok product demo, a customer's unboxing photo, or a livestreamed event, visuals shape how consumers perceive, evaluate, and engage with brands, often more than traditional, text-based communication (Sample et al., 2020). For brand managers, this creates a challenge: confronted with an enormous volume of complex visual data, traditional manual analysis often proves time-consuming and laborious, making it challenging to uncover subtle insights accurately.

Computer vision technology has become a powerful ally in brand management. By leveraging advanced machine learning, particularly deep learning, computer vision algorithms enable automated interpretation of images or videos, by recognizing elements such as color, texture, objects, and emotional expressions (Szeliski, 2011). For example, brands can monitor how their logos are used across user-generated content, detect emerging consumer associations, or even predict emotional responses to visual advertising. Beyond efficiency and accuracy, these insights can inform strategy from product design to campaign optimization.

In this article, we show how computer vision technologies can help strengthen brand identity, clarify brand meaning, measure consumer response, and deepen brand relationships. We also explore how integrating computer vision with other data sources—like text, audio, or video—can enhance personalization, enable real-time feedback loops, and support rebranding efforts. When brands learn to "see" as their consumers do, they gain a powerful competitive edge (Li et al., 2025).

How visual features drive brand value

From logos, packaging to influencer video and store design, visual elements are crucial to how brands communicate meaning and build equity. But not all visual content is created or processed in the same way. To make sense of the growing volume of image-based data, we developed a practical framework that breaks visual content into three levels: low-, high-, and integrative-level (Li et al., 2025).

- *Low-Level Visual Features:* basic and primitive visual aspects of an image, such as color (hue, brightness, and saturation), texture (coarseness, directionality), shape (corners, curves), and spatial composition (positioning, distance, and symmetry). These "low-level" features are fundamental building blocks in photography and brand communication (e.g., packaging design and layout of online or in-store displays).
- *High-Level Visual Features:* semantic and interpretive elements derived from low-level features, such as people (e.g., influencers, consumers), objects (like products, brand logos), scenes (usage scenarios, brand atmosphere), and even actions or emotions. They carry more profound narrative or emotional content—for example, a brand might show consumers smiling in real usage situations or highlight scenic backdrops to convey a sense of vacation.
- *Integrative-level Visual Features:* visual aesthetics, complexity, and similarity that aggregate low- and high-level features.

Recent research on computer vision in branding shows a clear shift from analyzing simple visual cues to examining more complex and meaningful elements. Managers can no longer rely solely on static images or basic descriptors if they want a full picture of how their brand is portrayed and perceived. Instead, they consider "integrative" visual attributes—like aesthetics, complexity, and similarity—which can reveal how multiple features interact to shape consumer engagement and brand equity. This broader perspective extends beyond images alone: logos, videos, product packaging, store layouts, and even digital "servicescapes" in apps or websites are now relevant sources of insight. Many firms and researchers are also adopting multimodal approaches, combining visual, textual, and audio data to give a holistic understanding of brand perception and strategy. By using these richer methods, brand leaders can more effectively identify what truly resonates with consumers and tailor their messaging, design, and overall brand experience for a more significant impact.

How computer vision powers brand strategy

Computer vision has become a powerful tool that elevates branding research by revealing new layers of visual insight. We summarize different tasks and applications in branding in Table 1. It starts with low-level feature extraction, where descriptors like color or text histograms and Harris corner detector quantify the basic elements of brand visuals, everything from a product's dominant color to the roundness of a logo's corners (Dew et al., 2022).

Table 1. Computer vision tasks and applications in branding

Computer vision tasks	Description	Application in branding
Low-level feature extraction	Identifying specific visual characteristics (e.g., color, texture, shape) of images.	Copycat detection; identity and design consistency evaluation.
Objects detection	Identifying specific objects and their location in an image, such as cars, people, animals, or fashion items.	Celebrity detection and brand communication effect; brand-related object richness, diversity, placement, and engagement; logo detection and brand visibility tracking.
Image tagging/captioning	Providing a readable description of an image in keywords or sentences	User-generated images analysis and brand association; travel and leisure-related objects and scenes recognition for destination image.
Image classification	Categorizing an image into one of several pre-defined classes or categories.	Brand selfie type classification; brand personality classification; image quality classification or sentiment classification; consumer social profiling.
Image clustering	Grouping similar images by visual characteristics.	Similarity between influencers and brands; clustering of gastronomic photos and destination images; product or service recommendation.
Image segmentation	Dividing an image into multiple distinct regions or segments, each of which corresponds to a different object or part of the image.	Destination landscape depiction; destination perceptions and intentions; recognition of consumption and usage situation.

Face detection and facial expression recognition	Identifying human faces in images or videos and measuring human emotions or expressions through the detection and analysis of facial features.	Consumer sentiment analysis by facial expression recognition; consumer demographics analysis; evaluation of face trustworthiness and attractiveness of sellers.
Image generation	Using algorithms to generate new images that change style.	Enhanced packaging or product design; style transferring of brand logos; generation of logo.
Pose detection	Detecting people and localizing their key points, such as hands, joints, elbows, knees, and wrists.	Detection of hand gestures, pose, and body language of speakers in brand-related videos and images.
Movement detection	Tracking human movement direction and magnitude by identifying patterns of apparent motion.	Detection of motion direction and magnitude of speakers, products, or logos in brand-related videos.
Text from image	Extracting printed and handwritten text in supported languages from images.	Extraction of text information from brand-related images.
Prediction	Prediction based on visual features.	Prediction of brand image aesthetics, project success, brand-related social media post virality, consumer ratings, and review helpfulness.

From there, tasks such as object detection using CNN (convolutional neural network) and ViT (vision transformer) identify key brand elements in an image—logos, products, and even celebrities—and pinpoint where they appear. This deeper level of analysis helps marketers map out brand salience, detect copycats, and understand how product placement or fashion items might boost brand popularity. Similarly, image segmentation sketches elements (e.g., products, logos, settings) by dividing an image into multiple distinct regions or segments, enabling brand managers to fine-tune layouts and understand usage situations.

Beyond mere detection, image tagging and captioning algorithms assign descriptive labels or full-sentence captions to visuals, making it easier to gauge how brands are represented on social media or how users engage with certain brand elements. Meanwhile, image classification and clustering group together pictures that share themes or patterns, particularly useful for sorting user-generated content into categories like “packshots” or “brand selfies” (Hartmann et al., 2021).

Facial Expression Recognition captures viewers’ emotional cues (e.g., smiles, frowns), allowing managers to quickly assess how audiences respond to brand content and adjust strategies for greater resonance and impact. In a similar spirit, tasks like pose or movement detection can capture the dynamic aspects of consumer interaction, such as how people handle or showcase products in videos, offering glimpses into real-time brand engagement (Zhou et al., 2021).

The technology also paves the way for image generation, where models like GANs (generative adversarial networks) and VAEs (variational autoencoders) can produce entirely new creative assets—fresh logos, innovative package designs, or product concepts (Dew et al., 2022). Research in this domain is still in its early stage, but the potential for personalization and rapid prototyping is significant.

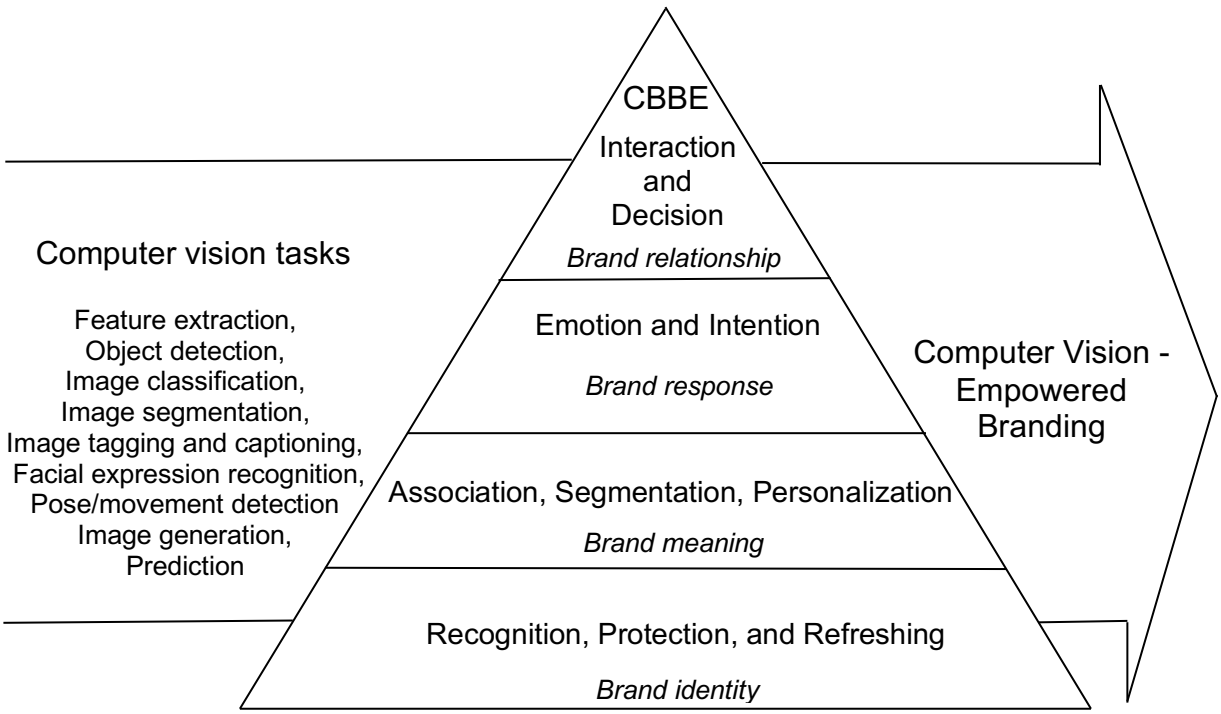
Finally, prediction with visual input enhances forecasting and decision-making. By incorporating visual data into traditional analytic models, researchers have become markedly more successful at predicting marketing outcomes from brand post virality to business survival rates (Zhang & Luo, 2022).

Despite the wealth of insights these tasks provide, computer vision's capacity to interpret deeper cultural, emotional, or symbolic meanings remains limited. That is why many researchers combine automated approaches with qualitative methods to enrich the findings. This ensures that while computers handle large-scale analyses, the nuances of human perception and brand storytelling are never lost.

Empowering brand strategy with computer vision

We delineate four approaches to empower brand strategy with computer vision based on the consumer-based brand equity framework (CBBE, Keller 1993). Figure 1 depicts the four approaches in detail. CBBE illustrates the brand equity framework as a pyramid, where brand identity lays the foundation, brand meaning stretches the abstract image, brand response captures consumers' reaction, and brand relationship emphasizes the deep connection between consumer and brand.

Figure 1. Framework for empowering brand strategy with computer vision



Approach 1: Reinforcing brand identity: recognition, protection, and refreshing

Managers face the challenge of ensuring a brand's visual identity remains consistent across diverse markets, platforms, and retail environments. Inconsistent use of logos, colors, or packaging can dilute brand recognition and customer trust. Computer vision offers a scalable solution by automatically verifying whether brand visuals appear in their correct format, color, and shape across touchpoints (Dew et al., 2022). This enables marketers to maintain brand coherence even on a global scale. Strategically, such monitoring safeguards brand equity and ensures a uniform consumer experience.

Another pressing concern for managers is the threat of counterfeit or look-alike products, which can erode market share and damage brand credibility. Computer vision systems can detect visual similarities, such as suspicious shapes, logos, or color schemes, and flag potential copycats in real time. This allows for early intervention before consumer confusion or reputational harm occurs. As a result, brands can more effectively police their visual territory and protect their market position.

Managers aiming to refresh or evolve their brand face the delicate task of innovating without alienating loyal customers. Computer vision and generative AI tools can support this process by generating visual variations in logos, packaging, or product design. These AI-generated prototypes help marketers test consumer reactions quickly through focus groups or digital platforms. While final creative decisions still rest with humans, these tools streamline the ideation process and help strike a balance between novelty and brand familiarity, thereby facilitating more strategic brand evolution.

Approach 2. Deepening brand meaning: association, segmentation, and personalization

Managers today face the challenge of keeping up with how their brand is portrayed not just by official channels, but by consumers themselves, through countless photos and videos shared online. These user-generated visuals often reflect real-world associations, usage contexts, and aspirational meanings that consumers attribute to the brand. Computer vision allows marketers to analyze these images at scale, uncovering organic patterns such as popular product pairings or lifestyle signals (Liu et al., 2020). Strategically, these insights can inform brand storytelling and help managers emphasize the most resonant attributes in future campaigns.

In a world of increasingly personalized experiences, managers must find ways to tailor messages without over-relying on traditional demographics. Computer vision can identify features such as age, gender, or emotional expressions within consumer-shared visuals and pair them with existing CRM or e-commerce data. This enables precise segmentation and hyper-targeted marketing. As a result, brands can offer more relevant product recommendations and communications, enhancing both satisfaction and loyalty.

Approach 3. Understanding consumer response: emotion and intention

Managers often struggle to capture authentic, real-time emotional responses to their brand experiences, whether during live events, retail interactions, or social media engagement. Traditional feedback tools may miss subtle cues or produce results too late to be actionable. Computer vision, through face detection and emotion-recognition algorithms, allows brands to assess consumers' emotional states as expressed in selfies, group photos, and videos. This enables real-time insight into whether customers appear happy, neutral, or dissatisfied (Lu et al., 2016). Strategically, such feedback can inform immediate adjustments, like tweaking lighting or interactivity at a pop-up store, or validate the appeal of a new product before a larger launch.

As visual word-of-mouth plays a growing role in shaping consumer decisions, managers need to understand how photo-based reviews and user-generated content contribute to brand attitude and purchase intention. Computer vision can analyze these visuals to detect aesthetic quality, experiential elements, and how well text and image content align (Ceylan et al., 2024). These cues are critical in influencing consumer intention. With these insights,

brands can fine-tune content guidelines and encourage more impactful user submissions, ultimately strengthening their visual word-of-mouth strategy.

Approach 4. Nurturing brand relationships: interaction and decision

Brand managers often face the challenge of understanding what drives consumer loyalty, engagement, and advocacy in digital spaces. Likes, shares, and comments on branded images are valuable signals—but identifying what specifically fuels them can be elusive. Computer vision enables marketers to analyze visual features such as colorfulness, human presence, and facial expressions to pinpoint which styles spark the most engagement (Li & Xie, 2020). By uncovering patterns linked to specific audience preferences, brands can fine-tune their visual strategy to deepen emotional connections and amplify brand equity across social media platforms.

Another challenge lies in connecting branding efforts with concrete performance metrics like conversions, bookings, or sales. In platforms like Airbnb, e-commerce, or food delivery services, the visual quality of listings plays a crucial role in consumer decision-making. Computer vision can assess image quality, composition, and content, linking these features to downstream behavior, such as click-through rates, booking frequency, or product purchases (Zhang et al., 2022). Strategically, this empowers managers to forecast demand more accurately, optimize pricing, and improve visual standards to drive business success.

Conclusion and recommendations

For brand managers, computer vision is no longer a distant frontier; it is a highly actionable tool for improving brand consistency, gaining real-time consumer insight, and shaping emotional engagement at scale. As technology matures, the question is not whether to adopt computer vision, but how to do so effectively and responsibly. Below are key considerations for marketers looking to integrate computer vision into their branding strategy:

1. Clarify your strategic goals

Before implementation, define what you want to achieve with computer vision: Is it visual consistency across touchpoints? Improved consumer segmentation? More engaging content? A higher conversion rate? Different goals require different computer vision capabilities, from logo detection and emotion recognition to aesthetic scoring and predictive modeling.

2. Blend automation with human creativity

While computer vision excels at large-scale pattern detection and rapid visual analysis, it cannot replace human intuition, cultural awareness, or storytelling. Use computer vision to guide creative direction, not dictate it. The most successful brands will combine algorithmic efficiency with creative finesse to preserve the brand's soul.

3. Ensure ethical and privacy-conscious use

When analyzing user-generated images or employing facial recognition, privacy and consent become paramount. Establish clear guidelines and robust data protection protocols to avoid reputational risks and ensure compliance with regional laws (e.g., GDPR).

4. Focus on cross-functional collaboration

Computer vision deployment isn't just a marketing task—it involves data teams, legal advisors, IT, and creative departments. Cultivating collaboration across functions will ensure smoother integration, better outcomes, and stronger buy-in across the organization.

5. Embrace multimodal and longitudinal insights

Don't treat computer vision insights in isolation. Combine image analysis with text, audio, or self-report data to build richer consumer profiles. Over time, cross-platform and longitudinal analysis can reveal how brand visuals shape (and are shaped by) evolving consumer sentiment and loyalty.

6. Aim for smart targeting and personalization

Combining predictive analytics with computer vision methods enables advanced profiling, personalized pricing, and targeted promotions, ultimately enriching consumer experiences and reinforcing brand strength in an increasingly complex digital landscape.

7. Leverage generative AI for co-creation

Generative tools, powered by computer vision, can accelerate design iteration and support consumer co-creation. Marketers might invite consumers to contribute ideas for logos, packaging, or campaign visuals, then assess submissions with computer vision tools to ensure alignment with brand identity and visual coherence.

In an era saturated with content, computer vision gives marketers a powerful lens for detecting what truly resonates. When used strategically and ethically, it enables brands not only to listen and learn but also to adapt and co-create brand value together with consumers.

References

Ceylan, G., Diehl, K., & Proserpio, D. (2024). Words meet photos: When and why photos increase review helpfulness. Journal *of Marketing Research, 61*(1), 5–26.

Dew, R., Ansari, A., & Toubia, O. (2022). Letting logos speak: Leveraging multiview representation learning for data-driven branding and logo design. *Marketing Science, 41*(2), 401–425.

Hartmann, J., Heitmann, M., Schamp, C., & Netzer, O. (2021). The power of brand selfies. *Journal of Marketing Research, 58*(6), 1159–1177.

Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing, 57*(1), 1–22.

Li, Y., Lee, H. H. M., & Blasco-Arcas, L. (2025). Computer vision in branding: A conceptual framework and future research agenda. *Journal of Business Research*, 193, 115329.

Li, Y., & Xie, Y. (2020). Is a picture worth a thousand words? An empirical study of image content and social media engagement. *Journal of Marketing Research, 57*(1), 1–19.

Liu, L., Dzyabura, D., & Mizik, N. (2020). Visual listening in: Extracting brand image portrayed on social media. *Marketing Science, 39*(4), 669–686.

Lu, S., Xiao, L., & Ding, M. (2016). A video-based automated recommender (VAR) system for garments. *Marketing Science, 35*(3), 484–510.

Sample, K. L., Hagtvedt, H., & Brasel, S. A. (2020). Components of visual perception in marketing contexts: A conceptual framework and review. *Journal of the Academy of Marketing Science, 48*(3), 405–421.

Szelisk, R. (2011). Computer vision: Algorithms applications. Springer-Verlag.

Wedel, M., & Pieters, R. (Eds.). (2007). *Visual marketing: From attention to action*. Psychology Press.

Zhang, M., & Luo, L. (2022). Can consumer-posted photos serve as a leading indicator of restaurant survival? Evidence from Yelp. *Management Science, 69*(1), 25–50.

Zhang, S., Lee, D., Singh, P. V., & Srinivasan, K. (2022). What makes a good image? Airbnb demand analytics leveraging interpretable image features. *Management Science, 68*(8), 5644–5666.

Zhou, M., Chen, G. H., Ferreira, P., & Smith, M. D. (2021). Consumer behavior in the online classroom: Using video analytics and machine learning to understand the consumption of video courseware. *Journal of Marketing Research, 58*(6), 1079–1100.