

Sesgos cognitivos en la publicidad digital de alimentos y su impacto emocional: un enfoque comparativo basado en modelos de inteligencia artificial y visión por computadora

Cognitive biases in digital food advertising and their emotional impact: a comparative approach based on artificial intelligence and computer vision models

Recibido: 16 de septiembre de 2025

Aprobado: 15 de diciembre de 2025

Forma de citar: K. Morales and K. Y. Sánchez Mojica, "Sesgos cognitivos en la publicidad digital de alimentos y su impacto emocional: un enfoque comparativo basado en modelos de inteligencia artificial y visión por computadora", Mundo Fesc, vol. 16, no. 34, pp. 292-312 Jan. 2026, doi: 10.61799/2216-0388.2093.

Katherine Morales Gutiérrez 

Ingeniera en formación
kmorales34@unisalle.edu.co
ORCID: 0009-0005-1631-3612
Universidad de La Salle
Itagüí, Antioquia, Colombia

Karla Yohana Sánchez-Mojica 

Doctora en Educación
karla.sanchez@unisalle.edu.co
ORCID: 0000-0003-3164-4725
Universidad de La Salle
Cúcuta, Norte de Santander, Colombia

*Autor para correspondencia:
kmorales34@unisalle.edu.co



Sesgos cognitivos en la publicidad digital de alimentos y su impacto emocional: un enfoque comparativo basado en modelos de inteligencia artificial y visión por computadora

Resumen

La publicidad digital de alimentos explota sistemáticamente los sesgos visuales perceptivos para influir tanto en las emociones como en el comportamiento de los consumidores; sin embargo, la investigación existente carece de marcos computacionales que vinculen la detección automatizada de los sesgos visuales con la inferencia emocional probabilística. El objetivo de este artículo es diseñar y validar una arquitectura híbrida para inferir perfiles emocionales probables basados en los sesgos visuales de las imágenes publicitarias de alimentos, es decir, una combinación de técnicas de visión artificial profunda y modelización probabilística. Con este fin, se construyó una arquitectura de dos módulos: el primero se basa en un modelo de visión profunda con clasificación multietiqueta, utilizando el modelo EfficientNet-B0, que se comparó con los modelos ResNet-50, ViT-B/16 y MobileNet-V3 Large; el segundo fue el componente probabilístico correspondiente, basado en una cadena de Markov que asigna las muestras de sesgos detectadas a diez categorías emocionales relevantes para el marketing alimentario. También se creó un conjunto de datos compuesto por 1375 imágenes de publicidad alimentaria procedentes de las redes sociales, que se anotaron manualmente en once categorías establecidas de sesgo visual. En cuanto a los resultados, la arquitectura EfficientNet-B0 obtuvo las siguientes puntuaciones: precisión media = 0,961, AUC del ROC = 0,968 y puntuación F1 ponderada = 0,898, superando a las arquitecturas implementadas en este estudio, mientras que el módulo afectivo inspirado en Markov generó perfiles emocionales que se interpretan fácilmente en la literatura empírica. En definitiva, la conclusión es que la arquitectura propuesta separa el nivel perceptivo del nivel afectivo, lo que permite una explicabilidad mecánica, actualizaciones a través de módulos y la extensibilidad del espacio emocional, además de facilitar su aplicación práctica en la investigación de mercados, la salud pública y la regulación de la comunicación comercial.

Palabras clave: cadena de Markov, clasificación multietiqueta, modelado afectivo, publicidad alimentaria, sesgos visuales, visión por computadora.

Cognitive biases in digital food advertising and their emotional impact: a comparative approach based on artificial intelligence and computer vision models

Abstract

Digital food advertising systematically exploits perceptual visual biases to influence both emotions and consumer behavior; however, the existing body of research lacks computational frameworks that link the automated detection of visual biases with probabilistic emotional inference. The purpose of this article is to design and validate a hybrid architecture for inferring probable emotional profiles based on visual biases in food advertising images, that is, a combination of Deep Computer Vision techniques and probabilistic modeling. To this end, a two-module architecture was constructed: the first is based on a Deep Computer Vision model with multi-label classification, using the EfficientNet-B0 model, which was compared with the ResNet-50, ViT-B/16, and MobileNet-V3 Large models; the second was the corresponding probabilistic component based on a Markov chain that maps the detected bias samples into ten emotional categories relevant to food marketing. A dataset was also created comprising 1,375 food advertising images from social media, which were manually annotated across eleven established categories of visual bias. Regarding the results, the EfficientNet-B0 architecture achieved the following scores: mean Average Precision = 0.961, ROC-AUC = 0.968, and weighted F1 score = 0.898, outperforming the architectures implemented in this study, while the Markov-inspired affective module produced emotional profiles that are easily interpreted in empirical literature. Finally, the conclusion is that the proposed architecture decouples the perceptual level from the affective level, enabling mechanistic explainability, updates via modules, and the extensibility of the emotional space, and allowing for practical applicability in marketing research, public health, and the regulation of commercial communication.

Keywords: affective modeling, computer vision, food advertising, Markov chain, multilabel classification, visual biases.

Introducción

Currently, exposure to food advertising through recommendation algorithms is increasingly common [1], [2]. This phenomenon not only encourages the consumption of certain products but also contributes to the formation of unhealthy eating habits and is associated with increased obesity and related diseases, especially in populations with fewer economic resources [3]–[5]. Perceptual psychology and visual advertising research have established how the elements of framing, color, font and the placement of brands within visual elements influence the interpretation of messages via perceptual bias. Perceptual bias refers to the systematic and relatively unthinking way cognitive processing occurs with respect to directing our attention [6], [7]. These perceptual biases represent an ongoing, predictable pattern of behavior that exists in visuals and have been proven to be reliable predictors of both purchase intention and actual product selection at the time of sale [8]–[10].

However, significant methodological gaps remain. Many studies analyze the relationship between image and purchase decision using limited stimuli and self-report measures, making it difficult to extrapolate these findings to real-world digital advertising environments [11], [12]. Similarly, computer vision models applied to food advertising tend to focus on visual content without considering the persuasive mechanisms that structure the advertising scene, and they rarely compare architectures in multi-label classification contexts [13], [14]. Furthermore, few approaches quantitatively link the presence of visual biases to specific emotional categories using reproducible probabilistic models [14], [15], which limit comparability between studies and their usefulness in regulatory contexts [3]. To address these limitations, this study proposes the development and evaluation of an artificial intelligence architecture aimed at the automatic detection of visual biases in food advertising and the inference of probabilistic emotional profiles. To this end, a comparative approach is implemented that evaluates four deep learning architectures integrated with a probabilistic model based on transition matrices that allows estimating the relationship between detected visual biases and associated emotional states [13], [16], [17] [18].

Visual biases in food advertising images

An image is considered biased when it induces emotions or judgments about a product beyond what is explicitly communicated [19]. In food advertising, these mechanisms are used to highlight desirable product attributes and minimize less attractive characteristics, such as the degree of processing or high caloric density [7], [20]. Various visual resources influence this process, including framing, colors, typography, brand placement, and gaze direction, which guide attention and promote product recall and purchase intention [10], [21]–[25]. Based on these principles, the literature has identified multiple relevant visual biases [2], [7], [18], [26]–[31]. However, most of these resources have been studied through manual coding and isolated analysis, which limits the understanding of how multiple biases interact in a single advertisement [32], [33]. Furthermore, although around sixteen relevant visual biases have been identified, several categories present conceptual overlaps, which hinders their direct implementation in annotation schemes and computational models [7], [8].

As a result, the initial 16 theoretical biases were reorganized into 11 aggregated categories, which constitute the final classes of the model: visual_grouping_bias, visual_direction_bias, Framing bias, visual salience, and visual distance, brand bias, visual_metaphor_bias, movement_bias, sharpness_and_overlap bias, cognitive proximity bias, texture_bias, visual_typography_bias, aesthetic_stylistic_and_symmetry_bias [34]–[36]. Based on the grouping of key visual resources in food advertising [13], [32].

Bias classification models analyzed

The selection of computer vision models is based on representational capacity, computational efficiency, and transfer learning. EfficientNet-B0 stands out for its composite scaling and its ability to capture visual and semantic features relevant to detecting bias in advertising images [35], [37]. ResNet-50 provides representational depth through residual connections, establishing a robust baseline for multiple classification [38]. Vision Transformer (ViT-B/16) allows for modeling global dependencies through attentional mechanisms, facilitating the identification of complex visual configurations [39]. Finally, MobileNet-V3 Large is included for its computational efficiency in resource-constrained environments [40], [41].

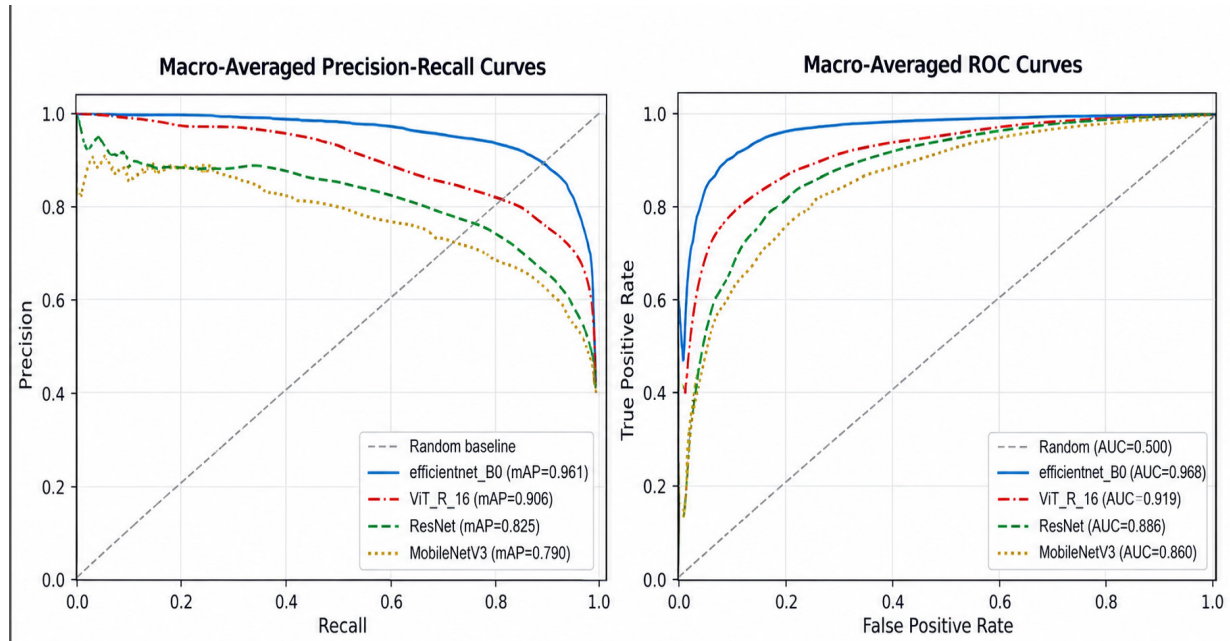
The architectures were evaluated by transfer learning and fine-tuning with food advertising images (train = 1100, test = 275).

TABLE I. Comparison of global multilabel classification metrics for the four architectures evaluated. The model with the best performance in each metric is highlighted in blue.

MODEL PERFORMANCE OVERVIEW & COMPARATIVE ANALYSIS									
Model	mAP	PR-AUC	ROC-AUC	LRAP	F1 micro	F1 macro	F1 ponder.	FPS (Inf.)	Latency (ms)
EfficientNet-B0	0.961 ↑	0.961 ↑	0.968 ↑	0.927 ↑	0.900 ↑	0.892 ↑	0.898 ↑	35	39
ViT-B/16	0.906	0.906	0.919	0.882	0.829	0.815	0.821	15	12
ResNet-50	0.825	0.825	0.886	0.904	0.764	0.755	0.792	60	23
YOLOv8	0.895	0.895	0.908	0.865	0.810	0.795	0.801	35	22
Swin Transformer	0.912	0.912	0.925	0.888	0.835	0.821	0.827	55	10

Table I and Fig. 1 present the comparison of global multilabel classification metrics for the four architectures evaluated under the metrics (F1-macro, mAP and LRAP) [42]. EfficientNet-B0 performed best across all metrics, achieving an mAP and PR-AUC of 0.961, a ROC-AUC of 0.968, and a weighted F1 of 0.898, positioning it as the most robust model for classifying food advertising images.

Fig. 1. Averaged precision-recall curves and averaged ROC curves.



Since multi-label detection reliably quantifies the presence of biases in an image, its outputs can feed a probabilistic module estimating $P(\text{Emotion}|\text{Bias})$, enabling the emotional profiling of any advertisement.

Computational model for the analysis of visual persuasion

The literature on affective image analysis shows that visual stimuli can be associated with emotional states using computational models [43]. In food advertising, several studies have identified visual properties linked to pleasure, desire, or rejection; however, most model the image-emotion relationship without explicitly considering the role of visual biases. This limitation makes it difficult to attribute emotional effects to specific configurations when multiple biases coexist in the same image. To overcome this limitation, an additional transitive level is introduced that estimates, for each image, a bias probability vector $P(s_i | \text{imagen})$, where s_i corresponds to one of the 11 defined visual biases. At the affective level, a transition matrix $P(\text{Emoción} | \text{Sesgo})$ represents empirical knowledge about the association between each bias and emotional responses. Interpreted as a Markov chain over the emotional space conditioned by the biases [15], this structure allows us to infer, using the law of total probability, the emotional distribution induced by each advertisement:

$$P(e_j | \text{imagen}) = \sum_{i=1}^{11} P(e_j | s_i) \cdot P(s_i | \text{imagen}) \quad (\text{Eq.1})$$

The transition matrix is constructed from quantitative empirical evidence documenting the stochastic effect of bias in a set of individuals and integrates both basic emotions and affective states specific to the food context. Probability estimates for every element $P(e_j | s_i)$ consist of effect size weighted counts which have been normalized so that they form a valid probability distribution across every row. Because there are some bias-emotion combinations that do not have enough or varying amounts of evidence, Laplace smoothing is used to produce estimate based on these conditions, mathematically defined as:

$$P(e_j | s_i) = \frac{c_{ij} + \alpha}{\sum_k (c_{ik} + \alpha)} \quad (\text{Eq. 2})$$

where c_{ij} is the weighted total amount of the relationship between a bias (s_i) and emotion (e_j) and $\alpha > 0$ is a regularization constant that guarantees non-negative probabilities across all possible combinations, maintains the standard definition of a transition matrix (row sums are equal to one), and enables the creation of a Markov chain in case of sparsity with respect to bias or emotion in the literature [65][66] Table II presents this correspondence, which defines the model's emotional vocabulary and grounds the computation of the $P(\text{Emotion}|\text{Bias})$ matrix.

TABLE II. Visual biases associated with associated emotional responses and effect indices reported in the literature.

Visual/Perceptual bias	Emotional Response	Effect Index	Metrics	Academic Source
1. Visual Grouping	Aesthetic pleasure	70–85%	Estimated	P1[44]
	Purchase desire / Motivation	65–75	%Inferred	P2 [45]
2. Visual Direction	Curiosity / Anticipation	70%+	Pdirect measurement3	[11], [46]
	Tension / Uncertainty	40–55%	Inferred P4	[47]
3. Framing / Visual Saliency				

Fear / Aversion (loss)	72–78%	P direct measurement 5	[48]
Hope / Desire (gain)	65–72%	Direct measurement P 6	[6]
Amygdala activation (urgency)	$r = 0.80^{***}$	fMRI correlation 7	[6]
4. Brand Bias			
Trust / Loyalty	$= 0.39; R = 0.71$	Regression ($p = .001$)8	[49]
Nostalgia / Identity belonging	60–75%	Estimated P 9	[50]
5. Visual Metaphor			
Surprise / Cognitive Pleasure	65–80%	Estimated P 10	[19]
Curiosity / Engagement	55–70%	Estimated P (P300)11	[51]
6. Motion Bias			
Excitement / Arousal	70–85%	Direct measurement P 12	[52]
Urgency / Tension	55–65%	Estimated P 13	[53]
7. Visual Sharpness and Overlap			
Confidence / Clarity / Pleasure (high sharpness)	65–80%	Estimated P 14	[54], [44]
Confusion / Anxiety (high overlap)	50–65%	Estimated P 15	[55]
8. Cognitive Proximity (Mere Exposure)			
Familiarity / Comfort / Positive Affect	$r = 0.26$ (208 studies)	Correlation (meta-analysis) 16	[50], [56]
Belonging / Security (family brand)	60–75%	Inferred P 17	[57], [45]
9. Visual Texture			
Tactile desire / Sensory pleasure (organic textures)	55–70%	Estimated P 18	[58]
Coldness / Efficiency (smooth textures)	50–65%	Estimated P 19	[59]
10. Visual Typography			
Trust / Sincerity (humanist serif)	+12–13%	positive % ($n = 400$) [20,c]	[60]
Clarity / Modernity (clean sans-serif)	+5–8%	positive Estimated % [21,c]	[60]
11. Symmetry / Aesthetics			
Aesthetic pleasure / Admiration	75–90%	Direct measurement P 22	[43]
Serenity / Balance	65–75%	Estimated P 23	
Fascination / Awe (moderate complexity)	55–65%	Estimated P 24	[62] [61]

The probabilities reported in Table II as estimated P were not extracted directly from the cited original studies but derived by approximate conversion of the effect sizes reported in those studies, following the conceptual frameworks of [63], [64].

It is important to note that these conversions are heuristic approximations and not exact calculations derived from the original data. Their purpose is to facilitate qualitative comparison between effects reported using heterogeneous metrics. They should not be interpreted as empirical probabilities directly measured in the source studies.[67] [68] The proposed matrix thus articulates a computational and probabilistic framework linking visual biases with emotional states.

Methodology

This study uses a mixed-methods approach combining both quantitative and qualitative data to develop a computational inference model of emotional profiles created by visual biases in food advertisements. The system has two modules connected. The first is a computer vision model that calculates the probability of different visual biases being present in an image. The second is a probabilistic model that predicts the emotional space of these visual biases based on the Markov transition matrix.

2.1. Downloading and cleaning the data

The image database was built from food advertising campaigns available online, specifically ads designed for social media. English queries were formulated that combined food products with visual compositions typical of digital advertising (banners, close-ups, prominent textures, blurred backgrounds, high color contrast, among others). The collection was performed automatically using the GoogleImageCrawler class from the icrawler library, limiting the number of images to a maximum of 1,500 per composition group and a maximum number of downloads per term.

2.2. Manual labeling

Visual biases were modeled as a multilabel classification problem, allowing each image to belong simultaneously to several categories. Each resulting image was represented as a binary vector of length 11, where each component indicates the presence (1) or absence (0) of the consolidated biases.

2.3. Model Training

The architecture of our model is a multi-label convolutional neural network (CNN) built in PyTorch to detect bias. We used a pre-trained EfficientNet-B0 as the base model on ImageNet, replacing its classifier with a dropout layer ($p = 0.5$) and creating 11 output neurons that operate as logit values. At inference, we applied a sigmoid activation function only to generate per-class probabilities.

All images were resized to 224×224 using random resized crops and normalized using the statistics from ImageNet (torchvision.transforms.v2). Our training data was augmented to create greater compositional and chromatic variation while retaining the semantics of bias. A partial fine-tuning strategy was employed. First, we used EfficientNet-B0 as a feature extractor, keeping all its weights frozen, and then unfreezing the last set of blocks and the new classifier to balance domain adaptation with overfitting mitigation in the constrained data environment. Our dataset was divided into 80% training/20% testing (retaining the multilabel distribution). During training, we used a batch size of 128, and trained for a maximum of 40 epochs, stopping early with a tolerance of six epochs on the validation loss. All training was run on a GPU using mixed precision, autocasting, and gradient scaling.

Maximizing performance via loss function optimization. Cross entropy with logits was used for the loss function, which had an associated positional weight determined by the training frequency for each label to give more weight to labels with lower frequencies (also known as unbalanced multilabel classification). The optimizer used for this project was the AdamW optimizer; there were two separate parameter groups for the classifier (learning rate of 3×10^{-4}) and for the unfreeze efficientnet blocks (learning rate of 5×10^{-4}). A weight decay of 1×10^{-4} was configured for both groups. The learning rate schedule used was a cosine annealing (max epochs in the Tmax field). To help with overfitting several regularization techniques were applied, including dropouts, weight decay and gradient clipping. The final evaluation of model performance was based on the lowest validation loss checkpoint.

For evaluating the model's performance, since this is multilabel classification, the in-training metric was micro F1 score on validation data (from trained model passed through sigmoid thresholded logits). The testing metric was a global threshold (which was established by analyzing the precision/recall curve from the validation results), set to 0.60. Final evaluations included reporting the per-label, macro and weighted F1's, as well as precision, recall and true and false positive and negative counts for each label. The assessments of ranking quality were conducted based on per-label average precision scores; mean AP and label ranking average precision scores. The selected model was selected because it provided the best compromise between mean average precision (mean AP) and macro F1 score.



Fig. 2. Diagram of the proposed model.

2.4. Emotional modeling using the matrix P(Emoción | Sesgo)

To construct the Markov matrix, Laplace smoothing was applied with a parameter $\alpha = 0.1$, chosen to balance two opposing risks. On the one hand, too low a value would have maintained the problematic influence of zeros in the original matrix (emotions not attributed to bias due to a lack of documentation or study), limiting the model's ability to capture emotional transitions poorly documented in the literature. On the other hand, too high a value would have diluted the differences between bias-emotion pairs with strong evidence and those without empirical support, treating both cases as equally probable. The value of 0.1 allows for the introduction of a minimum probability in empty cells without altering the hierarchy established by empirical data among documented transitions, thus giving greater weight to available evidence and less to unobserved phenomena.

Results

Results

3.1. Classified Images

The manual classification carried out to train the bias recognition model allowed us to obtain a set of 1500 images labeled and segmented into the different categories of the study. This is an initial proposal that contributes to the analysis of biases using computational methods and to carry out a more in-depth study in the field of emotional analysis.

3.2. Proposed architecture and results of the final model training

The training performance indicates that the model can be run without the explicit need for long training sessions or large datasets, as seen in Fig. 3, each of the categories presents high accuracy, in which the visual typography bias and the cognitive proximity bias stand out.

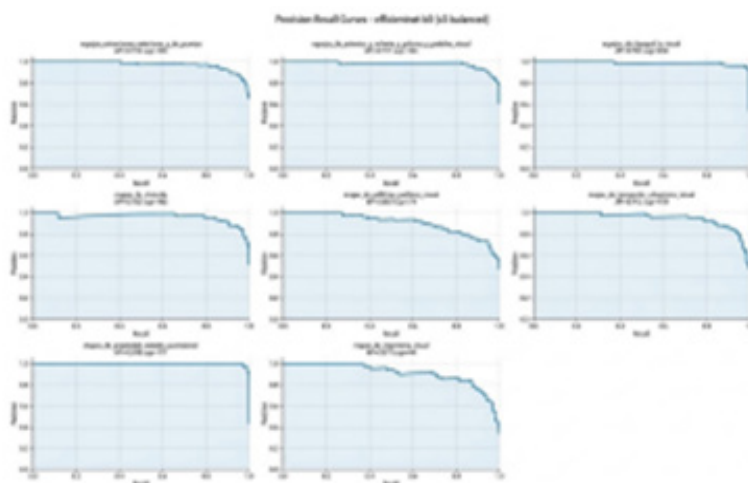


Fig. 3. Accuracy-recall curves by class for modified EfficientNet-B0.

Visual Typography and Cognitive Proximity were the most accurately classified categories, because both categories are based on features that are highly salient to EfficientNet-B0 (text, font styles, the presence of people or animals). In contrast, Sharpness and Visual Directional Proximity had the lowest classification accuracies, because the cues that define these categories (a blurred background, directional indicators) co-occur with framing and cognitive-separation proximity in food ads; therefore, the high degree of overlap between these features biases the predictions towards the more frequent category.

3.3. Results of Laplace smoothing

Visual Typography and Cognitive Proximity were the most accurately classified categories, because both categories are based on features that are highly salient to EfficientNet-B0 (text, font styles, the presence of people or animals). In contrast, Sharpness and Visual Directional Proximity had the lowest classification accuracies, because the cues that define these categories (a blurred background, directional indicators) co-occur with framing and cognitive-separation proximity in food ads; therefore, the high degree of overlap between these features biases the predictions towards the more frequent category.

3.3. Results of Laplace smoothing

TABLE III. Normalized probability emotional model (Markovian matrix)

Visual Bias (S _i)	E ₁ Pleasure (aesthetic)	E ₂ Desire / Motivation	E ₃ Curiosity / Antic.	E ₄ Trust / Security	E ₅ Fear / Aversion	E ₆ Excitement / Arousal	E ₇ Surprise / Cognitive pleasure	E ₈ Familiarity / Nostalgia	E ₉ Serenity / Balance	E ₁₀ Confusion / Anxiety
S ₁ Visual Grouping	0.3535	0.3232	0.0404	0.0404	0.0404	0.0404	0.0404	0.0404	0.0404	0.0404
S ₂ Visual Direction	0.0460	0.0460	0.3678	0.0460	0.2644	0.0460	0.0460	0.0460	0.0460	0.0460
S ₃ Framing / Visual Salience	0.0324	0.2545	0.0324	0.0324	0.2755	0.2431	0.0324	0.0324	0.0324	0.0324
S ₄ Brand Bias	0.0419	0.0419	0.0419	0.3396	0.0419	0.0419	0.0419	0.3249	0.0419	0.0419
S ₅ Visual Metaphor	0.0426	0.0426	0.3085	0.0426	0.0426	0.0426	0.3511	0.0426	0.0426	0.0426
S ₆ Motion Bias	0.0421	0.0421	0.0421	0.0421	0.2947	0.3684	0.0421	0.0421	0.0421	0.0421
S ₇ Sharpness / Visual Overlap	0.0435	0.0435	0.0435	0.3587	0.0435	0.0435	0.0435	0.0435	0.0435	0.2935
S ₈ Cognitive Proximity (Mere Exposure)	0.0434	0.0434	0.0434	0.3362	0.0434	0.0434	0.0434	0.3167	0.0434	0.0434
S ₉ Visual Texture	0.0455	0.3295	0.0455	0.0455	0.0455	0.0455	0.0455	0.0455	0.3068	0.0455
S ₁₀ Visual Typography	0.0437	0.0437	0.0437	0.3603	0.0437	0.0437	0.0437	0.0437	0.2904	0.0437
S ₁₁ Symmetry / Aesthetics	0.2960	0.0320	0.0320	0.0320	0.0320	0.0320	0.2240	0.0320	0.2560	0.0320

Discussion

The findings indicate that a computational image bias emotional model for food advertising is achievable. Complex visual patterns generated by EfficientNet B0 (including text, framing, color, and aesthetic characteristics) were successfully recovered through a restricted training set. Performance similarity of four different architectures (ViT B/16; ResNet 50; MobileNet V3) indicates the systemic nature of these visual biases and that these patterns will translate across all architectures. This ability to independently replicate visual bias detection through multiple deep learning methods demonstrates that automated detection should be considered a legitimate method of market research and regulatory assessment.

The proposed pipeline addresses three separate deficiencies that exist in the current literature. First, visual biases are addressed as a coherent integrated set rather than distinct individual components like most studies treat them. Second, images are used exclusively as inputs into the model; therefore, there is no requirement to use self-reported data as part of the analysis. Finally, visual biases are linked to individual emotional states using a probabilistic graphical model that supports the capability of making quantitative comparisons between studies [3].

A key limitation of the study sample was that it consisted of a sample set drawn from within Western nations, which may limit our ability to generalize these findings to other cultures. Future research will need to test the model across different cultural contexts and to add more comprehensive definitions to the bias taxonomy.

Conclusions

Based on empirical data, 90% or more of the analysed food ads used in some way visual graphical representations that were recognizable to their audiences as being biased, often on several occasions in one image. Therefore, visual design as a persuasive tool uses recurring, familiar forms rather than irrationally based style examples. The EfficientNet-B0 architecture achieved 961 mAP, 968 ROC-AUC, and 898 weighted F1 scores using only 1100 training and 275 test images without extensive GPU resources. The performance of ViT-B/16, ResNet-50, and MobileNet-V3 is similar but not as good as that of EfficientNet. Additionally, the biases detected in these other models are structurally regular and largely independent of their architectural inductive biases, which enables groups with limited resources to replicate the study's methodology.

The most visually salient classes (text, typography, people, animals) received the best scores for classification accuracy. The results from two biases that are related to the sharpness and direction of proximity did not perform well because their visual features overlapped with the framing and cognitive-separation bias, thereby causing the model's predictions to err towards the majority class as represented in the training data. These findings suggest that there likely will be adjustments made to the existing taxonomy classification as opposed to any errors in these model architectures.

The pipeline separates the perceived emotion and affective state and connects them together through an explicit Markov transition matrix. $P(\text{Emotion}|\text{Bias})$ is estimated by effect-size-weighted counts on the existing literature while applying Laplace smoothing ($\alpha = 0.1$) to develop a coherent ten-category visual vocabulary of emotion while maintaining the hierarchical evidence structure present in the literature. The above decisions produce three outcomes: the perceptual portion of the pipeline can be modified independently from the affective section; as new data become available it will be possible to extend the affective space without having to retrain; and that a quantifiable, reproducible visual-to-affective connection exists thus helping to partially address the previously documented methodological gap in this area.

Currently, food advertising regulations focus on textual claims (health, nutrition, and age) and do not address the visual form of persuasion. The automated capabilities presented here allow for visual auditing of items and supporting evidence of visual persuasion that was not previously available to regulators. Regulators can now delineate and classify their advertisements by bias of their composition and deduce the emotive quality of them through automated visual auditing. This enables them to add a new form of evidence to their regulatory framework for products that affect the public's health.

References

- [1] E. Boyland et al., "Food marketing, eating and health outcomes in children and adults: A systematic review and meta-analysis," *British Journal of Nutrition*, vol. 133, no. 6, pp. 781–805, 2025, doi: <https://doi.org/10.1017/S0007114524000102>
- [2] K. Kim, S. Lee, and Y. K. Choi, "Image proximity in advertising appeals: Spatial distance and product types," *Journal of Business Research*, vol. 99, pp. 490–497, 2019, doi: <https://doi.org/10.1016/j.jbusres.2017.08.031>
- [3] J. R. Cummings, L. V. Hoover, and A. N. Gearhardt, "A randomized experiment of the effects of food advertisements on food-related emotional expectations in adults," *Journal of Health Psychology*, vol. 28, no. 10, pp. 929–942, 2023, doi: <https://doi.org/10.1177/13591053231168340>

- [4] C. Baquedano et al., "The Power of Food Advertisements: A Brief Mindfulness Instruction Does Not Prevent Psychophysiological Responses Triggered by Food Ads," *Brain Sciences*, vol. 15, no. 3, p. 240, 2025, doi: <https://doi.org/10.3390/brainsci15030240>
- [5] F. Barbosa Escobar, Q. J. Wang, A. Corredor, and C. Velasco, "The taste of visual textures," *Food Quality and Preference*, vol. 100, p. 104602, 2022, doi: <https://doi.org/10.1016/j.foodqual.2022.104602>
- [6] B. De Martino, D. Kumaran, B. Seymour, and R. J. Dolan, "Frames, Biases, and Rational Decision-Making in the Human Brain," *Science*, vol. 313, no. 5787, pp. 684–687, 2006, doi: <https://doi.org/10.1126/science.1128356>
- [7] S. Fabbriizzi, S. Papadopoulos, E. Ntoutsis, and I. Kompatsiaris, "A Survey on Bias in Visual Datasets," *arXiv:2107.07919*, 2022, doi: <https://doi.org/10.48550/arXiv.2107.07919>
- [8] J. Dai, J. Cone, and J. Moher, "Perceptual salience influences food choices independently of health and taste preferences," *Cognitive Research: Principles and Implications*, vol. 5, no. 1, p. 2, 2020, doi: <https://doi.org/10.1186/s41235-019-0203-2>
- [9] J. Liang, Y.-Q. Zou, S.-Y. Liang, Y.-W. Wu, and W.-J. Yan, "Emotional Gaze: The Effects of Gaze Direction on the Perception of Facial Emotions," *Frontiers in Psychology*, vol. 12, p. 684357, 2021, doi: <https://doi.org/10.3389/fpsyg.2021.684357>
- [10] L. Miccoli, R. Delgado, P. Guerra, F. Versace, S. Rodríguez-Ruiz, and M. C. Fernández-Santaella, "Affective Pictures and the Open Library of Affective Foods (OLAF): Tools to Investigate Emotions toward Food in Adults," *PLOS ONE*, vol. 11, no. 8, p. e0158991, 2016, doi: <https://doi.org/10.1371/journal.pone.0158991>
- [11] S. Adil, S. Lacoste-Badie, and O. Droulers, "Face Presence and Gaze Direction in Print Advertisements: How They Influence Consumer Responses—An Eye-Tracking Study," *Journal of Advertising Research*, vol. 58, no. 4, pp. 443–455, 2018, doi: <https://doi.org/10.2501/JAR-2018-004>
- [12] M. Mehta and C. Buntain, "Emotional Images: Assessing Emotions in Images and Potential Biases in Generative Models," *arXiv:2411.05985*, 2024, doi: <https://doi.org/10.48550/arXiv.2411.05985>
- [13] R. Ismail and Z. Yuan, "Food Ingredients Recognition through Multi-label Learning," 2022.
- [14] Y. Shu, H. Gao, Y. Wang, and Y. Wei, "Food Emotional Perception and Eating Willingness Under Different Lighting Colors: A Preliminary Study Based on Consumer Facial Expression Analysis," *Foods*, vol. 14, no. 19, p. 3440, 2025, doi: <https://doi.org/10.3390/>

- [15] P. Cipresso, F. Borghesi, and A. Chirico, "Affects affect affects: A Markov Chain," *Frontiers in Psychology*, vol. 14, p. 1162655, 2023, doi: <https://doi.org/10.3389/fpsyg.2023.1162655>
- [16] F. Bockting, S. T. Radev, and P.-C. Bürkner, "Simulation-based prior knowledge elicitation for parametric Bayesian models," *Scientific Reports*, vol. 14, no. 1, p. 17330, 2024, doi: <https://doi.org/10.1038/s41598-024-68090-7>
- [17] A. M. Stefan, D. Katsimpokis, Q. F. Gronau, and E.-J. Wagenmakers, "Expert agreement in prior elicitation and its effects on Bayesian inference," *Psychonomic Bulletin & Review*, vol. 29, no. 5, pp. 1776–1794, 2022, doi: <https://doi.org/10.3758/s13423-022-02074-4>
- [18] S. Peker, G. G. Menekse Dalveren, and Y. _nal, "The Effects of the Content Elements of Online Banner Ads on Visual Attention: Evidence from An Eye-Tracking Study," *Future Internet*, vol. 13, no. 1, p. 18, 2021, doi: <https://doi.org/10.3390/fi13010018>
- [19] J. García-Madariaga, I. Moya, N. Recuero, and M.-F. Blasco, "Revealing Unconscious Consumer Reactions to Advertisements That Include Visual Metaphors. A Neurophysiological Experiment," *Frontiers in Psychology*, vol. 11, p. 760, 2020, doi: <https://doi.org/10.3389/fpsyg.2020.00760>
- [20] G. Hu et al., "Using artificial intelligence and computer vision to detect and monitor unhealthy child-directed food and beverage marketing: A data-driven study," *eClinicalMedicine*, vol. 89, p. 103549, 2025, doi: <https://doi.org/10.1016/j.eclinm.2025.103549>
- [21] I. Vermeir and G. Roose, "Visual Design Cues Impacting Food Choice: A Review and Future Research Agenda," *Foods*, vol. 9, no. 10, p. 1495, 2020, doi: <https://doi.org/10.3390/foods9101495>
- [22] A. Fenko, R. De Vries, and T. Van Rompay, "How Strong Is Your Coffee? The Influence of Visual Metaphors and Textual Claims on Consumers' Flavor Perception and Product Evaluation," *Frontiers in Psychology*, vol. 9, p. 53, 2018, doi: <https://doi.org/10.3389/fpsyg.2018.00053>
- [23] J. Palcu, J. Sudkamp, and A. Florack, "Judgments at Gaze Value: Gaze Cuing in Banner Advertisements, Its Effect on Attention Allocation and Product Judgments," *Frontiers in Psychology*, vol. 8, p. 881, 2017, doi: <https://doi.org/10.3389/fpsyg.2017.00881>
- [24] S. Spinelli, C. Masi, G. P. Zoboli, J. Prescott, and E. Monteleone, "Emotional responses to branded and unbranded foods," *Food Quality and Preference*, vol. 42, pp. 1–11, 2015, doi: <https://doi.org/10.1016/j.foodqual.2014.12.009>

- [25] A. Toet, M. G. Van Schaik, D. Kaneko, and J. B. F. Van Erp, "Do food cinemagraphs evoke stronger appetitive responses than stills?," *International Journal of Food Design*, vol. 4, no. 1, pp. 63–83, 2019, doi: https://doi.org/10.1386/ijfd.4.1.63_1
- [26] Y. Huang, X. Xue, E. Spelke, L. Huang, W. Zheng, and K. Peng, "The aesthetic preference for symmetry dissociates from early-emerging attention to symmetry," *Scientific Reports*, vol. 8, no. 1, p. 6263, 2018, doi: <https://doi.org/10.1038/s41598-018-24558-x>
- [27] A. R. Tartaglini, W. K. Vong, and B. M. Lake, "A Developmentally-Inspired Examination of Shape versus Texture Bias in Machines," *arXiv:2202.08340*, 2022, doi: <https://doi.org/10.48550/arXiv.2202.08340>
- [28] Y. Lu, "An Investigation of Gestalt Principles in Advertising Design," *Arts, Culture and Language*, vol. 1, no. 4, 2024, doi: <https://doi.org/10.61173/z3jn0p62>
- [29] M. Almourad, E. Bataineh, M. Hussein, and Z. Wattar, "Strategic Placement of Branding Elements in Digital Marketing: Insights from Eye-Tracking Data," in *Proc. 27th Int. Conf. on Enterprise Information Systems*, 2025, pp. 417–423, doi: <https://doi.org/10.5220/0013281500003929>
- [30] T. L. Childers and J. Jass, "All Dressed Up With Something to Say: Effects of Typeface Semantic Associations on Brand Perceptions and Consumer Memory," n.d.
- [31] M. J. Madsen, "A Q-Method Study of Visual Metaphors in Advertising," n.d.
- [32] G. Chochlakis, G. Mahajan, S. Baruah, K. Burghardt, K. Lerman, and S. Narayanan, "Using Emotion Embeddings to Transfer Knowledge Between Emotions, Languages, and Annotation Formats," *arXiv:2211.00171*, 2023, doi: <https://doi.org/10.48550/arXiv.2211.00171>
- [33] F. J. Otamendi and D. L. Sutil Martín, "The Emotional Effectiveness of Advertisement," *Frontiers in Psychology*, vol. 11, p. 2088, 2020, doi: <https://doi.org/10.3389/fpsyg.2020.02088>
- [34] K. Han et al., "A Survey on Visual Transformer," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 87–110, 2023, doi: <https://doi.org/10.1109/TPAMI.2022.3152247>
- [35] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *arXiv:1905.11946*, 2020, doi: <https://doi.org/10.48550/arXiv.1905.11946>
- [36] X. Wu et al., "CTransCNN: Combining transformer and CNN in multilabel medical image classification," *Knowledge-Based Systems*, vol. 281, p. 111030, 2023, doi: <https://doi.org/10.1016/j.knosys.2023.111030>
- [37] S. Kornblith, J. Shlens, and Q. V. Le, "Do Better ImageNet Models Transfer Better?," in *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 2656–2666, doi: <https://doi.org/10.1109/CVPR.2019.00277>

- [38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778, doi: <https://doi.org/10.1109/CVPR.2016.90>
- [39] A. Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," arXiv:2010.11929, 2021, doi: <https://doi.org/10.48550/arXiv.2010.11929>
- [40] A. Howard et al., "Searching for MobileNetV3," in Proc. IEEE/CVF Int. Conf. on Computer Vision (ICCV), 2019, pp. 1314–1324, doi: <https://doi.org/10.1109/ICCV.2019.00140>
- [41] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 4510–4520, doi: <https://doi.org/10.1109/CVPR.2018.00474>
- [42] O. Rainio, J. Teuho, and R. Klén, "Evaluation metrics and statistical tests for machine learning," Scientific Reports, vol. 14, no. 1, p. 6086, 2024, doi: <https://doi.org/10.1038/s41598-024-56706-x>
- [43] G. Lee, S. Yi, and J. Lee, "A Study on Deep Learning Performances of Identifying Images' Emotion: Comparing Performances of Three Algorithms to Analyze Fashion Items," Applied Sciences, vol. 15, no. 6, p. 3318, 2025, doi: <https://doi.org/10.3390/app15063318>
- [44] R. Reber, N. Schwarz, and P. Winkielman, "Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience?," Personality and Social Psychology Review, vol. 8, no. 4, pp. 364–382, 2004, doi: https://doi.org/10.1207/s15327957pspr0804_3
- [45] C. Janiszewski, "Preattentive Mere Exposure Effects," Journal of Consumer Research, vol. 20, no. 3, p. 376, 1993, doi: <https://doi.org/10.1086/209356>
- [46] C. K. Friesen and A. Kingstone, "The eyes have it! Reflexive orienting is triggered by nonpredictive gaze," Psychonomic Bulletin & Review, vol. 5, no. 3, pp. 490–495, 1998, doi: <https://doi.org/10.3758/BF03208827>
- [47] J. K. Maner, M. T. Gailliot, D. A. Rouby, and S. L. Miller, "Can't take my eyes off you: Attentional adhesion to mates and rivals," Journal of Personality and Social Psychology, vol. 93, no. 3, pp. 389–401, 2007, doi: <https://doi.org/10.1037/0022-3514.93.3.389>
- [48] A. Tversky and D. Kahneman, "The Framing of Decisions and the Psychology of Choice," Science, vol. 211, no. 4481, pp. 453–458, 1981, doi: <https://doi.org/10.1126/science.7455683>

- [49] V. Allom et al., "Reflective and impulsive processes underlying saving behavior and the additional roles of self-control and habit," *Journal of Neuroscience, Psychology, and Economics*, vol. 11, no. 3, pp. 135–146, 2018, doi: <https://doi.org/10.1037/npe0000093>
- [50] R. F. Bornstein, "Exposure and Affect: Overview and Meta-Analysis of Research, 1968–1987," *Psychological Bulletin*, vol. 106, no. 2, pp. 265–289, 1989, doi: <https://doi.org/10.1037/0033-2909.106.2.265>
- [51] L. Pileliene and V. Grigaliunaite, "The effect of female celebrity spokesperson in FMCG advertising: Neuromarketing approach," *Journal of Consumer Marketing*, vol. 34, no. 3, pp. 202–213, 2017, doi: <https://doi.org/10.1108/JCM-02-2016-1723>
- [52] P. Vuilleumier, "How brains beware: Neural mechanisms of emotional attention," *Trends in Cognitive Sciences*, vol. 9, no. 12, pp. 585–594, 2005, doi: <https://doi.org/10.1016/j.tics.2005.10.011>
- [53] L. Pessoa and R. Adolphs, "Emotion processing and the amygdala: From a 'low road' to 'many roads' of evaluating biological significance," *Nature Reviews Neuroscience*, vol. 11, no. 11, pp. 773–782, 2010, doi: <https://doi.org/10.1038/nrn2920>
- [54] P. Locher, E. A. Krupinski, C. Mello-Thoms, and C. F. Nodine, "Visual interest in pictorial art during an aesthetic experience," n.d.
- [55] D. M. Oppenheimer, "The secret life of fluency," *Trends in Cognitive Sciences*, vol. 12, no. 6, pp. 237–241, 2008, doi: <https://doi.org/10.1016/j.tics.2008.02.014>
- [56] R. B. Zajonc, "Attitudinal effects of mere exposure," *Journal of Personality and Social Psychology*, vol. 9, no. 2, pt.2, pp. 1–27, 1968, doi: <https://doi.org/10.1037/h0025848>
- [57] R. F. Bornstein and P. R. D'Agostino, "The Attribution and Discounting of Perceptual Fluency: Preliminary Tests of a Perceptual Fluency/Attributional Model of the Mere Exposure Effect," *Social Cognition*, vol. 12, no. 2, pp. 103–128, 1994, doi: <https://doi.org/10.1521/soco.1994.12.2.103>
- [58] A. Krishna, "An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior," *Journal of Consumer Psychology*, vol. 22, no. 3, pp. 332–351, 2012, doi: <https://doi.org/10.1016/j.jcps.2011.08.003>
- [59] A. Krishna and N. Schwarz, "Sensory marketing, embodiment, and grounded cognition: A review and introduction," *Journal of Consumer Psychology*, vol. 24, no. 2, pp. 159–168, 2014, doi: <https://doi.org/10.1016/j.jcps.2013.12.006>

- [60] M. Boulanger and P. Garnham, "Typography matters: The neuroscience of type," Monotype, 2022. [Online]. Available: <https://www.monotype.com/typography-matters>
- [61] I. Schindler et al., "Measuring aesthetic emotions: A review of the literature and a new assessment tool," PLOS ONE, vol. 12, no. 6, p. e0178899, 2017, doi: <https://doi.org/10.1371/journal.pone.0178899>
- [62] P. P. L. Tinio and H. Leder, "Just how stable are stable aesthetic features? Symmetry, complexity, and the jaws of massive familiarization," Acta Psychologica, vol. 130, no. 3, pp. 241–250, 2009, doi: <https://doi.org/10.1016/j.actpsy.2009.01.001>
- [63] J. Cohen, Statistical power analysis for the behavioral sciences, 2nd ed. Hillsdale, NJ, USA: L. Erlbaum Associates, 1988.
- [64] J. Ruscio and T. Mullen, "Confidence Intervals for the Probability of Superiority Effect Size Measurement and the Area Under a Receiver Operating Characteristic Curve," Multivariate Behavioral Research, vol. 47, no. 2, pp. 201–223, 2012, doi: <https://doi.org/10.1080/00273171.2012.658329>
- [65] M. Bondaschi et al., "From Markov to Laplace: How Mamba In-Context Learns Markov Chains," arXiv:2502.10178, 2025, doi: <https://doi.org/10.48550/arXiv.2502.10178>
- [66] J. M. Arengas Acosta, M. Lopez Ramirez, and R. Guzman Cabrera, "Impacto del preprocesamiento en la clasificación automática de textos usando aprendizaje supervisado y Reuters 21578," Revista Colombiana de Tecnologías de Avanzada (RCTA), vol. 1, no. 43, pp. 110–118, 2024, doi: [10.24054/rcta.v1i43.2506](https://doi.org/10.24054/rcta.v1i43.2506).
- [67] L. Contreras-Manrique, R. de B. Contreras-Manrique, T. V. Ovalle-Lizcano, and J. F. Rincón-Leal, "Análisis descriptivo de rasgos de personalidad en relación con el género y la edad de los estudiantes de la Universidad de Pamplona en la sede Villa Rosario," Eco Matemático, vol. 11, no. 1, pp. 51–61, 2020, doi: [10.22463/17948231.2943](https://doi.org/10.22463/17948231.2943).
- [68] A. M. Machacado-Rojas and L. E. Aparicio-Pico, "Técnicas de inteligencia artificial aplicadas al análisis de imágenes diagnóstico," Eco Matemático, vol. 12, no. 2, pp. 100–111, 2021, doi: [10.22463/17948231.3237](https://doi.org/10.22463/17948231.3237).
- [69] C. A. Mejia Rodriguez, M. A. Rincon Pinzon, L. M. Palmera Quintero, and L. M. Arevalo Vergel, "Aplicación de machine learning y metodología CRISP-DM para la clasificación precisa de severidad en casos de dengue," Revista Colombiana de Tecnologías de Avanzada (RCTA), vol. 1, no. 43, pp. 78–85, 2024, doi: [10.24054/rcta.v1i43.2822](https://doi.org/10.24054/rcta.v1i43.2822).

- [68] A. M. Machacado-Rojas and L. E. Aparicio-Pico, "Técnicas de inteligencia artificial aplicadas al análisis de imágenes diagnóstico," *Eco Matemático*, vol. 12, no. 2, pp. 100–111, 2021, doi: 10.22463/17948231.3237.
- [69] C. A. Mejia Rodriguez, M. A. Rincon Pinzon, L. M. Palmera Quintero, and L. M. Arevalo Vergel, "Aplicación de machine learning y metodología CRISP-DM para la clasificación precisa de severidad en casos de dengue," *Revista Colombiana de Tecnologías de Avanzada (RCTA)*, vol. 1, no. 43, pp. 78–85, 2024, doi: 10.24054/rcta.v1i43.2822.