

# Advancing Visual Food Attractiveness Predictions for Healthy Food Recommender System

Ayoub El Majjodi<sup>1,\*</sup>, Sohail Ahmed Khan<sup>1</sup>, Alain D. Starke<sup>1,2</sup>, Mehdi Elahi<sup>1</sup> and Christoph Trattner<sup>1</sup>

<sup>1</sup>MediaFutures, University of Bergen, Lars Hilles Gate 30, Bergen, Norway

<sup>2</sup>Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Amsterdam, Netherlands

## Abstract

The visual representation of food on digital platforms affects the foods chosen by users, including in the context of recommender systems. Previous studies show that small changes in visual features can influence human decision-making, regardless of whether the food is healthy. This paper reports on a study aimed at better understanding how users perceive the attractiveness of food recipe images in the digital world. In an online mixed-methods survey ( $N = 192$ ), users provided visual attractiveness ratings of food images on a 7-point scale, along with textual assessments. We found robust correlations between fundamental visual features (e.g., contrast, colorfulness) and perceived image attractiveness. The analysis also revealed that, among other user factors, cooking skills positively affected perceived image attractiveness. Regarding food image dimensions, appearance and perceived healthiness were significantly correlated with user ratings of food image attractiveness.

## Keywords

Food recommender systems, User modeling, Image attractiveness, Health, Personalization, Digital nudges

## 1. Introduction

Visual cues and attractiveness play a crucial role in everyday food choices [1]. Even when only presented with a food image, humans tend to instantly assess a food's energy density, expected taste and other characteristics [2]. As such, images are one of the key determinants of food preferences [2, 3], tapping into emotional and hedonic processes of an individual [4].

The importance of visual attractiveness also applies to digital choice context, including food recommender systems [5]. Our previous research has shown the capability of recommender systems to influence food behaviors via visual features, including the promotion of either high-fat or low-fat food choices [6], as well as encouraging the search for healthier options [3]. Additionally, our earlier work has established that visual attractiveness significantly contributes to predicting the online popularity of food items [7], and these visual features can also be leveraged to infer cultural backgrounds [8].

What is currently missing is an in-depth examination of image feature modeling. Although previous studies have extracted image features and examined the relation between those features, visual attractiveness and user preferences [6, 3], these models have not been optimized. Moreover, to date, image features have not been related to user characteristics (e.g., demographics, food knowledge), which are also important determinants of food preferences [9].

We present the results of a mixed-method study that explores the determinants of visual attractiveness in digital recipe images more comprehensively. Our approach builds upon previous work by modeling perceived visual attractiveness based on low-level image features [10, 11, 3]. Additionally, we seek to

---

*HealthRecSys'24: The 6th Workshop on Health Recommender Systems co-located with ACM RecSys 2024*

\*Corresponding author.

✉ ayoub.majjodi@gmail.com (A. E. Majjodi); sohail.khan@uib.no (S. A. Khan); a.d.starke@uva.nl (A. D. Starke); mehdi.elahi@uib.no (M. Elahi); christoph.trattner@uib.no (C. Trattner)

🌐 <http://ayoubmajjodi.info/> (A. E. Majjodi); <https://mediafutures.no/2021/09/13/sohail-ahmed-khan/> (S. A. Khan);

<https://mediafutures.no/2021/06/21/alain-starke/> (A. D. Starke); <https://mediafutures.no/2020/11/25/mehdi-elahi/> (M. Elahi); <https://christophtrattner.info/> (C. Trattner)

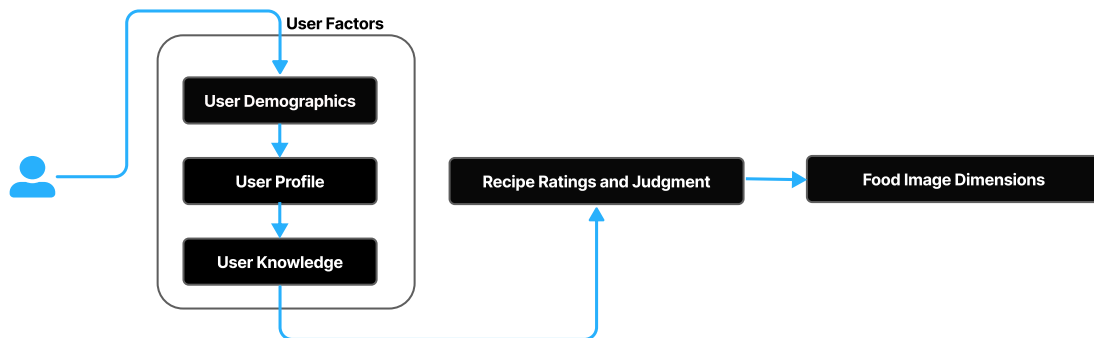
🆔 0000-0002-7478-5811 (A. E. Majjodi); 0000-0001-5351-2278 (S. A. Khan); 0000-0002-9873-8016 (A. D. Starke); 0000-0003-2203-9195 (M. Elahi); 0000-0002-1193-0508 (C. Trattner)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

optimize this model by integrating user characteristics that have been employed in knowledge-based food recommender systems to promote healthier recipe choices [12, 13, 14].

Finally, we inquire more qualitatively on user justifications for provided visual attractiveness ratings, asking to motivate their quantitative judgment. We formulate the following research questions:



**Figure 1:** Steps of the user flow designed for the online survey.

- **RQ1:** To what extent do the latest deep learning methods predict visual attractiveness compared to state-of-the-art low-level features?
- **RQ2:** To what extent do user characteristics, including demographics, food knowledge, and eating goals, predict food image attractiveness?
- **RQ3:** What dimensions determine the attractiveness of food image?

## 1.1. Contributions

Compared to our extensive previous work in the field of visual attractiveness and food recommender systems [6, 7, 8, 13], this study offers novel insights into several key aspects:

- Previous work mostly relied on low-level image attractiveness features, while this study shows how new deep-learning models compare to these old features.
- This work, compared to any before, also shows as to what extent demographic features play a role in predicting visual food attractiveness. To our knowledge, no other work has shown this before.
- Finally, this study tries to go beyond traditional quantitative black box approaches and reveals why images are rated less or more attractive.

## 2. Study Design

To perform our study, we employed a dataset sourced from the well-known recipes website All-Recipes.com, with the addition of new recipe photos [3, 14]. The dataset comprised various recipe features, including image URL, ingredients, amount of fats and sugar, and instructions and ingredients. To generate a diverse set of images, we randomly selected 200 recipes with relatively from the dataset of 58,000. As most images in this dataset were relatively unattractive [3], we used the recipe’s title in search engines and image websites (e.g., Unsplash) to look for more attractive images for 100 of these recipes. To validate this process, three computational food researchers, including a co-author, voted on which of the two photos was the most attractive to ensure a diverse set of recipe images in terms of expected attractiveness.

The study involved a survey design, as depicted in Figure 1. Participants first provided demographic information, as well as responded to items that measured their subjective food knowledge (4 items) and cooking skills (6 items), using 5-point Likert scales based on earlier work [15, 16, 17]. We also used questions from earlier work on a knowledge-based food recommender [14], to inquire on other

user characteristics, including recipe website usage and home cooking frequency, cooking experience and dietary goals. Afterwards, users were invited to rate the visual attractiveness of 12 semi-randomly selected recipe images, on 7-point attractiveness scales. In addition, to address [RQ3], they were asked to write at least one sentence about why they had given this rating. Finally, to support our examination of [RQ3], we used 5-point Likert scales on food image dimensions [18], to ask to what extent a recipe’s appearance, expected taste, healthiness, and familiarity affected their attractiveness ratings.

We employed the Prolific crowdsourcing platform to recruit 192 users (65% male;  $M_{age} = 33.54$ ) to participate in our study. The study took approximately 11 min to complete and participants were reimbursed with GBP 1.65<sup>1</sup>.

**Table 1**

Linear regression models predicting visual attractiveness ratings for recipe images: (A) with low-level image visual features, (B) with deep learning-based visual features. \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .

(A)		(B)			
Low-level Image Features		Image features Extractor			
	$\beta$ (S.E)	VGG16	ResNet	CLIP	
Colourfulness	6.725 (1.521)***				
Brightness	2.136 (0.155)***				
Naturalness	1.925 (0.530)***				
Entropy	1.026 (0.154)***				
Saturation	-3.976 (1.020)***	$R^2$	0.351***	0.349***	0.357***
Sharpness	-1.182 (1.187)*	RMSE	1.500	1.491	1.501
RGBContrast	-1.782 (3.808)				
Contrast	7.401 (11.101)				
Constant	-6.884 (1.243)***				
$R^2$	0.110***				
RMSE	1.753				

### 3. Results

To address our research questions, we primarily employed linear regression models. This helped to understand the principal impacts of image attributes and user characteristics on image attractiveness, the latter derived from user ratings. For our thematic analysis, the images were split into attractive and unattractive based on the mid-point of the rating scale (4) ( $M = 4.33$ ,  $SD = 1.80$ ). Details of used materials and conducted analyses can be accessed through the following URL [19].

#### 3.1. RQ1: Predicting Visual Attractiveness

We first modeled perceived visual attractiveness based on the underlying image features. We extracted diverse low-level visual features using the OpenIMAJ Java Framework (cf. [7]). Subsequently, we conducted a linear regression analysis to predict attractiveness based on these extracted visual features. The results are outlined in Table (1.A), revealing that several image features significantly affected the attractiveness of a recipe image:  $F(8, 2100) = 32.66$ ,  $p < 0.001$ . Specifically, Colourfulness, Brightness, Naturalness, and Entropy demonstrated a positive association with image attractiveness. In contrast, Saturation, Sharpness, and RgbContrast negatively affected image attractiveness. In line with [3], these results suggested that users perceived colorful, bright, and naturalistic food images as more attractive.

<sup>1</sup>Our study complied with the ethical guidelines of the Research Council of Norway and the guidelines of University of Bergen for scientific research. It was judged to pass without further extensive review.

Going beyond low-level visual image features, we used deep learning architecture models. Our toolkit included established models, such as VGG16 [20] and ResNet [21], along with the well-known transformer [22] architecture for visual feature extraction, CLIP<sup>2</sup> [23]. Table (1.B) outlines the performance of these different models, outperforming our regression model in terms of  $R^2$  and RMSE. This aligns with previous research where deep learning embeddings also outperformed low-level visual features within the context of food [7, 24].

### 3.2. RQ2: User characteristics and Image Attractiveness

We further examined whether user factors affected the perceived visual attractiveness of images. Accordingly, we divided user characteristics into different categories: User demographics, User profile, which represented the backbone of a food knowledge-based recommender system, and User knowledge, which measures the user’s food knowledge and cooking skills. A confirmatory factor analysis, reported in Table 2, showed that both subjective food knowledge and cooking skills adhered to internal consistency guidelines ( $\alpha > 0.70$ ) while they also met the guidelines for convergent validity ( $AVE > 0.5$ ).

**Table 2**

Results of the principal component factor analysis across different subjective food knowledge and cooking skills. Items were measured on 5-point Likert scales. Cronbach’s Alpha is denoted by  $\alpha$ ,  $AVE$  is the average variance explained. Items in grey and without loading were omitted.

Aspect	Item	Loading
Subjective Food Knowledge $\alpha = 0.866$ $AVE = 0.858$	Compared with an average person, I know a lot about healthy eating.	0.777
	I think I know enough about healthy eating to feel pretty confident when choosing a recipe.	0.885
	I know a lot about how to evaluate the healthiness of a recipe.	0.773
	I do not feel very knowledgeable about healthy eating.	0.932
Cooking skills $\alpha = 0.783$ $AVE = 0.591$	I can confidently cook recipes with basic ingredients.	0.751
	I can confidently follow all the steps of simple recipes.	
	I can confidently taste new foods.	0.737
	I can confidently cook new foods and try new recipes.	0.869
	I enjoy cooking food.	0.655
	I am satisfied with my cooking skills.	0.816

Table (3.A) presents the outcomes of the linear regression model aimed at forecasting the attractiveness of image recipes:  $F(9, 2090) = 3.60$ . Among the various user factors examined, only two significantly affected recipe attractiveness: cooking skills ( $\beta = 0.34$ , p-value= 0.00021) and recipe website usage ( $\beta = 0.18$ , p-value= 0.020). However, none of the other user aspects affected user ratings for a given image recipe. Additionally, we also analyzed a combined model of image features and user factors, but this led to results similar to the separate models reported in Tables (1 and 3.A). This suggested that low-level visual features had a more significant impact on food image attractiveness than user features, largely in line with preliminary findings in previous research [3, 18].

### 3.3. RQ3: Justifications for Visual Attractiveness

To assess the influence of different food image dimensions on user ratings for food images, we modeled visual attractiveness based on the reported importance of food image dimensions. Table 3 outlines the results of the regression model:  $F(4, 21) = 2.41$ .

Two factors significantly impacted attractiveness. First, appearance significantly impacted user ratings ( $\beta = 0.12$ ,  $p = 0.03$ ). Second, the expected healthiness from the images also demonstrated a

<sup>2</sup>Contrastive Language-Image Pre-training (CLIP).



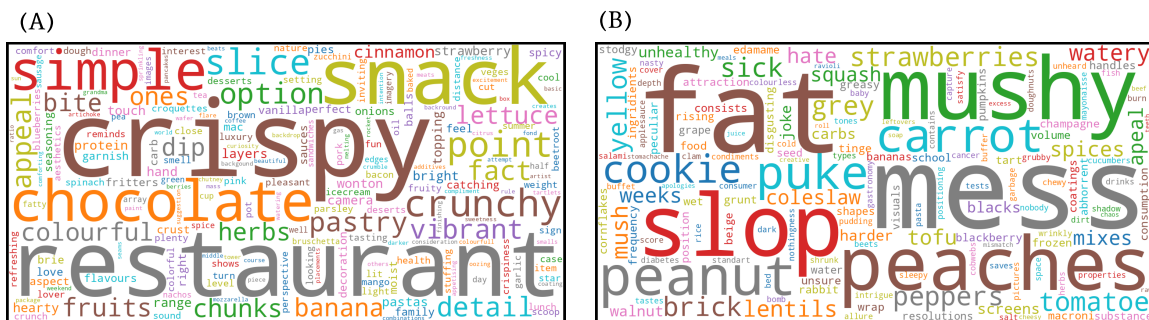
**Table 3**

Linear regression models predicting user rating for recipe image attractiveness: (A): with user factors, (B): with food image dimensions. \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .

(A)		(B)	
User Factors	$\beta$ (S.E)	Food Image Dimension	$\beta$ (S.E)
<b>User Demographic</b>			
Age	-0.047 (0.116)	Appearance	0.129 (0.061)*
Education	-0.424 (0.320)	Healthiness	0.077 (0.035)*
Gender	-0.077 (0.088)	Taste	-0.005 (0.050)
<b>User Profile</b>		Familiarity	0.0231 (0.038)
Recipe Website Usage	0.201 (0.086)*	Constant	3.487 (0.365)***
Home Cooking	-0.009 (0.078)	R <sup>2</sup>	0.011***
Cooking Experience	-0.052 (0.079)	RMSE	1.855
Eating Goals	0.019 (0.063)		
<b>User Knowledge</b>			
Subjective Food Knowledge	-0.213 (0.138)		
Cooking Skills	0.315 (0.086)***		
Constant	4.001 (0.570)***		
R <sup>2</sup>	0.015***		
RMSE	1.845		

significant impact ( $\beta = 0.07$ ,  $p = 0.03$ ). However, perceived taste and familiarity did not impact user ratings.

To gain insights into the reasons behind the visual attractiveness ratings, we collected qualitative justifications from users. Natural Language Processing (NLP) techniques, such as punctuation removal, repeated character elimination, and stop word filtering, were applied to analyze 2,019 user justifications for both attractive and unattractive images. From these responses, we generated two word clouds to highlight the most prominent terms.



**Figure 2:** Word cloud for terms in the user judgment: (A) : judgments for attractive images, (B): judgments for unattractive images.

Figure 2 presents the most frequent responses associated with both attractive and unattractive images. These findings are discussed in relation to the themes of ‘appearance’ and ‘health’ (cf. Table (3.B)).

### 3.3.1. Appearance-based justifications

Figure 4 shows a few examples of user textual justifications. Several participants, including user ( $U_a$ ), expressed the term ‘crispy’ in their assessments of attractive images, mainly referring to appearance. The word ‘simple’ is frequently used by users, such as user ( $U_b$ ), to convey the simplicity of recipe

content. In contrast, ‘mess’ was more commonly associated with judgments of unattractive food images, indicating their unappealing appearance. Moreover, the repeated use of the term ‘fat’ suggested that fatty foods were generally perceived as unattractive, as in judgments by users ( $U_c - d$ ).



( $U_a$ ): “looks juicy with nice crispy bits, which is nice and clear in the picture”



( $U_b$ ): “Interesting, slightly unusual, and does look visually appealing with simple ingredients presented well”



( $U_c$ ): “It looks messy and unappealing”



( $U_d$ ): “Too much carbs/fat”

**Figure 3:** Examples of images used in the study, associated with users’ textual judgments related to the appearance. ( $U_a$ ) and ( $U_b$ ) are textual justifications for attractive images, while ( $U_c$ ) and ( $U_d$ ) are justifications for unattractive images.

### 3.3.2. Healthiness-based justifications

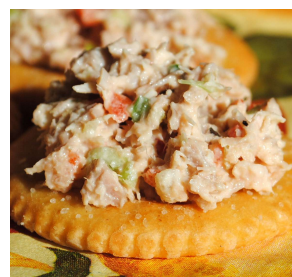
Judgments related to health frequently appeared in connection with the food’s appearance, such as by user ( $U_e$ ) in Figure ???. The term ‘restaurant’ was employed in various user judgments, often associated with presentation and healthiness, as described by the user ( $U_f$ ). Conversely, the concept of unhealthiness was linked to fatty foods and messy representation, as evident in the judgments of users ( $U_{g-h}$ ) in Figure ???.



( $U_a$ ): “Healthy salad option with balanced nutrients. It’s also quite colorful”



( $U_b$ ): “The dish looks very nice, like in a restaurant. It is colorful and looks very healthy”



( $U_c$ ): “It looks a bit mushy and brown and I don’t like Turkey”



( $U_d$ ): “Chicken is unhealthy and gross”

**Figure 4:** Examples of images used in the study, associated with users’ textual judgments related to the appearance. ( $U_a$ ) and ( $U_b$ ) are textual justifications for attractive images, while ( $U_c$ ) and ( $U_d$ ) are justifications for unattractive images.

## 4. Conclusion & Future work

This work has explored different aspects of the relationship between the user and food images. Through an online user study, we have found that various visual features can predict the attractiveness of a given image (i.e. colorfulness, brightness, naturalness). This prediction accuracy could be slightly improved using image features extracted using deep learning techniques (RQ1). In line with earlier work

[11, 3, 18], this suggests that the visual attractiveness of food images can be enhanced by increasing their colorfulness, brightness, and naturalness, while decreasing other features, such as saturating and sharpness. Obviously, there may be tradeoffs between these features when altering them.

Regarding user characteristics, none of the user demographics are related to food image attractiveness. In contrast, using online recipe websites and cooking skills are positively associated with the attractiveness of food images (RQ2). More novel is our contribution on the user justifications, for which we have found image appearance and perceived healthiness to be important dimensions of visual attractiveness ratings (RQ3). It seems that attractiveness are related to the expect taste or hedonic food goals (e.g., 'crispy'), while unattractive images focused on poor presentation and disliked ingredients.

Our study offers valuable insights into techniques for image attractiveness selection for various goals and domains. In particular, these techniques can be leveraged to persuade or nudge users towards specific eating goals, such as health [3, 25]. We believe that leveraging the visual appeal of attractive images can address this issue. Our future studies will focus on designing image selection pipelines for the application of food recommender systems tailored to guide people toward healthy food choices without compromising the benefits of personalization. We aim to analyze and categorize the collected textual judgment through thematic analysis to build word dictionaries related to image dimensions. These dictionaries can then be used to train learning models, enabling the evaluation of food image attractiveness based on user textual inputs.

## Acknowledgments

This work was supported by industry partners and the Research Council of Norway with funding to MediaFutures: Research Centre for Responsible Media Technology and Innovation, through the centers for Research-based Innovation scheme, project number 309339.

The authors acknowledge the use of ChatGPT [26] for checking and correcting the grammar of this article. No new content was generated this way; only existing text was checked and, if needed, corrected.

## References

- [1] I. Vermeir, G. Roose, Visual design cues impacting food choice: A review and future research agenda, *Foods* 9 (2020) 1495.
- [2] C. Spence, K. Motoki, O. Petit, Factors influencing the visual deliciousness/eye-appeal of food, *Food Quality and Preference* 102 (2022) 104672.
- [3] A. D. Starke, M. C. Willemsen, C. Trattner, Nudging healthy choices in food search through visual attractiveness, *Frontiers in Artificial Intelligence* 4 (2021) 20.
- [4] R. Cadario, P. Chandon, Which healthy eating nudges work best? a meta-analysis of field experiments, *Marketing Science* 39 (2020) 465–486.
- [5] D. Elswailer, H. Hauptmann, C. Trattner, *Food Recommender Systems*, Springer US, New York, NY, 2022, pp. 871–925. URL: [https://doi.org/10.1007/978-1-0716-2197-4\\_23](https://doi.org/10.1007/978-1-0716-2197-4_23). doi:10.1007/978-1-0716-2197-4\_23.
- [6] D. Elswailer, C. Trattner, M. Harvey, Exploiting food choice biases for healthier recipe recommendation, in: *Proceedings of the 40th international acm sigir conference on research and development in information retrieval*, 2017, pp. 575–584.
- [7] C. Trattner, D. Moesslang, D. Elswailer, On the predictability of the popularity of online recipes, *EPJ Data Science* 7 (2018) 1–39.
- [8] Q. Zhang, D. Elswailer, C. Trattner, Visual cultural biases in food classification, *Foods* 9 (2020) 823.
- [9] B. Scheibehenne, L. Miesler, P. M. Todd, Fast and frugal food choices: Uncovering individual decision heuristics, *Appetite* 49 (2007) 578–589.

- [10] A. Khosla, A. Das Sarma, R. Hamid, What makes an image popular?, in: Proceedings of the 23rd international conference on World wide web, 2014, pp. 867–876.
- [11] J. San Pedro, S. Siersdorfer, Ranking and classifying attractiveness of photos in folksonomies, in: Proceedings of the 18th international conference on World wide web, 2009, pp. 771–780.
- [12] C. Musto, C. Trattner, A. Starke, G. Semeraro, Towards a knowledge-aware food recommender system exploiting holistic user models, in: Proceedings of the 28th ACM conference on user modeling, adaptation and personalization, ACM, New York, NY, USA, 2020, pp. 333–337.
- [13] A. D. Starke, C. Musto, A. Rapp, G. Semeraro, C. Trattner, “tell me why”: using natural language justifications in a recipe recommender system to support healthier food choices, *User Modeling and User-Adapted Interaction* (2023) 1–34.
- [14] A. El Majjodi, A. D. Starke, M. Elahi, C. Trattner, et al., The interplay between food knowledge, nudges, and preference elicitation methods determines the evaluation of a recipe recommender system, in: Proceedings of the 10th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems (IntRS 2023), 2023, pp. 1–18.
- [15] L. R. Flynn, R. E. Goldsmith, A short, reliable measure of subjective knowledge, *Journal of business research* 46 (1999) 57–66.
- [16] Z. Pieniak, J. Aertsens, W. Verbeke, Subjective and objective knowledge as determinants of organic vegetables consumption, *Food quality and preference* 21 (2010) 581–588.
- [17] N. Frans, Development of cooking skills questionnaire for EFNEP participants in Kansas, Ph.D. thesis, Kansas State University, 2017.
- [18] Q. Zhang, D. Elswailer, C. Trattner, Understanding and predicting cross-cultural food preferences with online recipe images, *Information Processing & Management* 60 (2023) 103443.
- [19] A. El Majjodi, S. A. Khan, A. D. Starke, M. Elahi, C. Trattner, Examining the visual attractiveness of digital recipe images: Material, 2024. URL: <https://github.com/ayoubGL/Health-RecSys-2024>.
- [20] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556 (2014).
- [21] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
- [22] A. Vaswani, N. M. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is All you Need, in: *Neural Information Processing Systems*, 2017. URL: <https://api.semanticscholar.org/CorpusID:13756489>.
- [23] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, et al., Learning transferable visual models from natural language supervision, in: *International conference on machine learning*, PMLR, 2021, pp. 8748–8763.
- [24] J.-j. Chen, C.-W. Ngo, T.-S. Chua, Cross-modal recipe retrieval with rich food attributes, in: *Proceedings of the 25th ACM international conference on Multimedia*, 2017, pp. 1771–1779.
- [25] L. Yang, C.-K. Hsieh, H. Yang, J. P. Pollak, N. Dell, S. Belongie, C. Cole, D. Estrin, Yum-me: a personalized nutrient-based meal recommender system, *ACM Transactions on Information Systems (TOIS)* 36 (2017) 1–31.
- [26] OpenAI, Chatgpt, 2024. URL: <https://openai.com/chatgpt/>, accessed: 2024-09-16.