Understanding Advertisements

Yi Feng[1], Chuanyi Li[1]*, and Vincent Ng[2]

[1]State Key Laboratory for Novel Software Technology, Nanjing University, China
[2]Human Language Technology Research Institute, University of Texas at Dallas, USA
fy@nju.edu.cn, lcy@nju.edu.cn, vince@hlt.utdallas.edu

Abstract

While AI systems are capable of reading texts and seeing images, they typically perceive surface information explicitly conveyed with limited abilities to comprehend hidden messages (e.g., a double-edged remark). We propose the novel task of *advertisement understanding*: given an advertisement, which can be a text, an image, or a video, the goal is to identify the persuasion strategies used and determine the (possibly hidden) messages conveyed. Efforts on this task could enhance machine comprehension capabilities, and provide users with increased situation awareness w.r.t. the advertised message and thus possibly enable mindful decision making. We believe that this task presents long-term challenges to AI researchers and that successful understanding of ads could bring machine understanding one important step closer to human understanding.

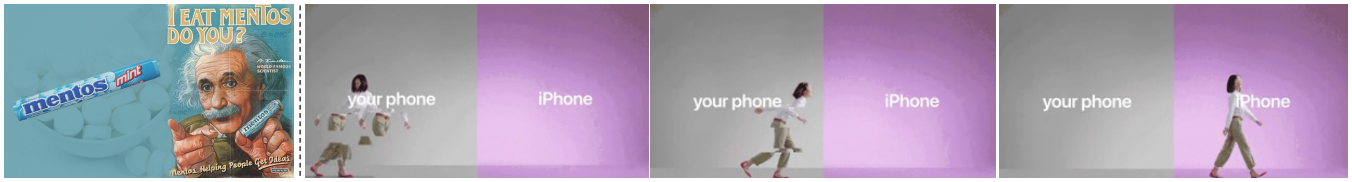
1 Introduction

With the advancement of technology, a new wave of digital media has emerged, including social media, short-form videos, and live streaming. These platforms enable individuals to share knowledge and express their opinions. Similarly, they provide enhanced opportunities for advertisers to promote their products and services. Different forms of advertisements (ads) have been permeating people's lives and bombarding people at all times. When opening an app, an ad often appears before one can access the main content. When watching YouTube, a video ad may unexpectedly interrupt the experience. Repeated exposure to these ads may influence viewers through psychological persuasion strategies, gradually steering them toward the advertisers' objectives. Often, viewers are unknowingly shaped by the underlying messages and intentions of the advertisers.

In recent years, the issue of manipulation in advertising has garnered increasing attention. While advertising is generally considered a legitimate commercial tool for promoting products fairly, some unethical advertisers exploit social media platforms such as Twitter and TikTok to deceive consumers and deliberately tarnish the reputations of competitors (Touahri and Mazroui 2024; Camara et al. 2024; Aberathne and Walgampaya 2021; Sun et al. 2021; Thejas et al. 2019; Blauth, Gstrein, and Zwitter 2022; Millar et al. 2018). Merchants often promote and exaggerate the quality of their products while selling inferior ones. Unfair business competitions result in a lose-lose situation for both parties. Meanwhile, advertising techniques such as user targeting and ad recommendations continue to evolve, allowing ads to be precisely targeted at consumers who are more likely to engage with them (Lin et al. 2024; Rasoolipour and Emamiifar 2020; Feng et al. 2021; Quan et al. 2020; Maio et al. 2021). Furthermore, there are also politicians who use advertising to create political manipulation during campaigns for election rigging (Horák et al. 2024; Wu et al. 2022b; Vijayaraghavan and Vosoughi 2022). Thus, there is a pressing need to alleviate ad manipulation and unveil the possibly hidden intentions behind ads.

Several persuasion strategies have been employed in ads to convey ideas and influence people. Persuasion aims at convincing the target to internalize the persuasive argument and adopt the new attitude (Rocklage, Rucker, and Nordgren 2018; Hogan 2010; Cialdini and Cialdini 2007), such as persuading consumers to buy products or watch the latest movies. There have been numerous persuasion strategies developed in advertising, from obvious ones to obscure and insidious ones (Andrews et al. 2013). One typical strategy is *Comparison*, where two competing products are compared to highlight the advantages of one over the other. Consider Figure 1(b), which is an animated image. The woman in the image moves from left to right, i.e., from "your phone" to "iPhone". During the movement, the woman's body becomes clearer and clearer. This strategy allows consumers to know that an iPhone runs smoothly and does not freeze, unlike other manufacturers' phones. The ultimate goal of this ad is to persuade consumers to buy iPhones, with the hidden message being *iPhones are better than other phones*. Many viewers would have no idea what the ad attempts to convey if they fail to notice the differences before and after the movement. Therefore, it is necessary to understand the persuasion strategies in order to fully comprehend the ad and the underlying hidden message. Note that the (explicitly expressed) surface message in this example is that a woman moves from left to right, while the (implicit) hidden message is *iPhones are better than other phones*. Thus, hidden messages are not the same as those that are explicitly described in ads. Moreover, hidden messages are different

*Corresponding author



(a) (b)

Figure 1: Examples of static and animated images. Selected frame tokens from the animated image are shown. The first example's hidden message is *Eating Mentos will make you as smart as Einstein*. The second example's hidden message is *iPhone is better than other phones* as the woman object in iPhone is clearer compared to that in phones from other manufacturers [1].

from ultimate goals, such as buying the products or clicking the links. They can be seen as claims established via a persuasive argumentation process that is eventually used to provide support for the ultimate goal. Nevertheless, some ads are straightforward and may not contain hidden messages (e.g., promotions on Amazon).

How can AI be used to understand ads in order to protect customers from being manipulated? We propose to use AI techniques to identify persuasion strategies and extract hidden messages conveyed in ads. Specifically, we define the novel task of *ad understanding*: given an ad, which can be a text, an image, or a video, the goal is to (1) detect the persuasion strategies, (2) locate where the strategies are, and (3) generate a description of the hidden messages, if any.

Given the prevalence of ads, it is somewhat surprising that little attention has been devoted to ad processing in the AI community (see Section 3 for a discussion of related work). From a research perspective, automated ad understanding presents a number of challenges to AI researchers:

Unveiling hidden messages General generation tasks, such as machine translation (Conia et al. 2024) and caption generation (Thakare and Walse 2022), typically follow a "what you see is what you get" approach, where the generated content directly reflects the expressed semantics or visual elements. However, unlike these tasks, ad understanding involves the challenging subtask of recovering the underlying arguments that advertisers deliberately withheld. Specifically, the key claim associated with an argument, along with the supporting evidences, are often intentionally made implicit, making it challenging to recover hidden messages. For instance, in Figure 1(a), current caption generation models can describe surface messages, such as *an elderly man eating candy*. In a more ideal scenario, the model might produce something like *Einstein eating Mentos*. In ad understanding, however, we need to infer the hidden message *Eating Mentos will make you as smart as Einstein*, which is beyond the capabilities of today's technologies. While the existing *argument mining* approaches developed in the natural language processing (NLP) community, which aim to identify the argumentative structure of text, can identify arguments that are *explicitly* expressed, they are unable to infer *hidden arguments* (i.e., arguments where the claims and/or premises are not explicitly expressed). A further complication is that ads typically feature minimal text with only a few words but are rich in visual elements. Extracting arguments from *visual elements* (as opposed to text) remains a little-explored area.

Injecting knowledge Background knowledge, especially historical and cultural knowledge, is often needed to properly understand ads. Consider Figure 1(a) again. Without the historical context that "Einstein is an intelligent scientist", it might be difficult to grasp the advertiser's intended message. Furthermore, semantic/visual shift poses a significant challenge, as individuals from diverse cultural backgrounds may interpret the same ad in different ways. For instance, a *green hat* refers to a hat that is green in color in western cultures, but this phrase can be used to signify that someone has been cheated on in some Asian cultures.

Understanding multiple modalities Ads may be represented in numerous modalities, such as text, image, video, and voice, and the hidden messages can be embedded in multiple modalities. For instance, to understand the ad in Figure 1(b), we need to capture the details in the visual modality (i.e., the woman is becoming increasingly clear from a blurred state) and combine them with extra information (e.g., rhetorical devices, such as the metaphor comparing the phone to the woman character). How to combine the information derived from different modalities to properly understand an ad remains an open question.

The rest of the paper is organized as follows. Section 2 describes the challenges involved in understanding different types of ads. Section 3 discusses related work. Sections 4 sketches the initial steps involved in automated ad understanding, including corpus construction and model design. Finally, we present our conclusions in Section 5.

2 Challenges

In this section, we discuss the challenges in processing different forms of ads.

Texts and Images

The most classic ads are image ads, which are usually accompanied by text. As shown in Figure 1(a), Einstein holds Mentos and tells everyone that he eats Mentos. In this example, the advertiser's message is *Eating Mentos will make you as smart as Einstein*.

[1] https://www.youtube.com/watch?v=kQcq3rpne78&t=15s&ab_channel=FahadThetics&t=15

Figure 2: An example of video ads[2]. Selected frame tokens from the video are shown. The hidden message is *Pepsi is better than Coke*.

When humans understand this ad, three steps are involved: (1) identifying the person and the sentences, (2) employing the historical knowledge that "Einstein is an intelligent scientist" and the commonsense knowledge that "Mentos is a mint", and (3) identifying the persuasion strategy and inferring the hidden message. Here, the persuasion strategy is *Testimonial*, which relies on having a celebrity support an idea or product as good or bad in order to influence people's opinions without letting them examine the facts carefully.

Compared with static images, animated images are more complicated. Consider Figure 1(b). This ad depicts a woman moving from left to right, having the hidden message *iPhones are better than other phones*. Besides *Comparison*, it employs a persuasion strategy known as *Name-Calling*, which involves linking a person or product with a negative symbol. Here, the vague woman in the ad is metaphorically linked to the frozen phone. The intent is to give the viewers the impression that other phones run slower than iPhones. Understanding this animated image requires that people take a further step to (1) pay more attention to the gesture and movement of the woman and (2) know the metaphor and associate the woman with any other phone.

Challenges For AI systems to identify persuasion strategies and extract hidden messages in text or image ads, they need to address the following issues:

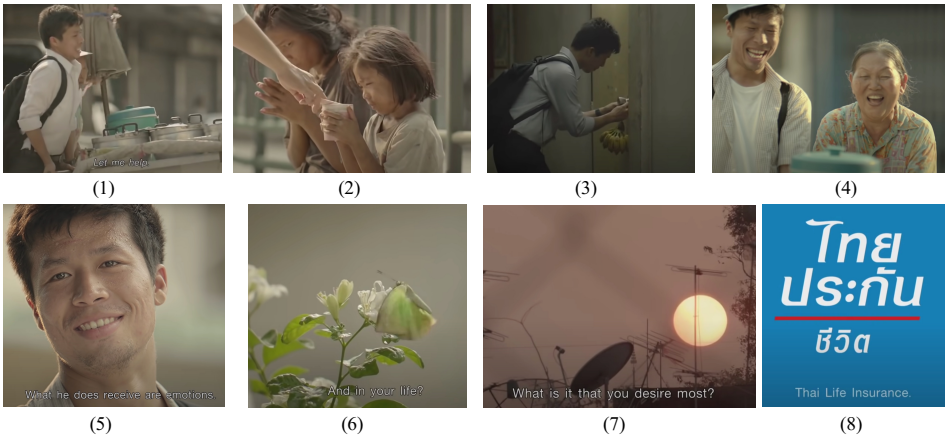
(1) Extracting detailed information in the ad, such as the words, clauses, sentences, and image objects. It is difficult to extract these details accurately from an ad. In spite of the possibility of using object detection (Zhou et al. 2021; Chen et al. 2021b) and named entity identification (Shen et al. 2022; Chen et al. 2022) techniques to extract words and image objects, these methods are from sufficient for achieving our goals. For instance, state-of-the-art object detectors are only effective at identifying large, common, sharp-edged objects, but are weak at detecting small objects with blurred boundaries in the background of images.

(2) Obtaining external knowledge that is not explicitly mentioned (e.g., Einstein is an intelligent scientist). Though there are several sources available for knowledge retrieval, such as the Internet and knowledge graphs (Chacon and Sosnovsky 2022), it is non-trivial to obtain such knowledge in practice. For example, even if AI systems successfully extracted the person from an image, how can they be certain that it was Einstein and not any other person? Further, even if AI systems are certain it was Einstein and subsequently searched for background information about Einstein on the Internet, how can they confirm which search results are accurate? While there exists work on exploring the injection of external knowledge into models (Yang et al. 2019; Seyler et al. 2018), the focus is on injecting a limited amount of domain-specific knowledge. Ad understanding, in contrast, requires a large amount of open-domain knowledge.

(3) Identifying the relationship between textual entities or image objects (e.g., by comparing the different objects corresponding to the woman in terms of location and clarity, we can infer that the iPhones run faster). Since there are multiple candidate woman objects in the example, how can AI systems select two proper candidates for comparison, and how can they capture the relationship between the selected woman objects? Image captioners (Fei 2022; Gao et al. 2022) could describe the general information of an image (e.g., a woman is walking), but they may neglect spatial details (e.g., a woman is walking from left to right and is becoming increasingly clear).

(4) Detecting rhetorical devices and their associations (e.g., the vague woman represents the frozen phone, which is the metaphorical device.). Numerous rhetorical devices exist (e.g., metaphor, sarcasm, hyperbole, idioms, and puns), and it is a challenge for AI systems to understand which rhetorical device is used in an ad. Moreover, how can AI systems know the metaphor associated with the detected device (e.g., linking a vague woman to a frozen phone)? While the majority of existing approaches to rhetorical device identification focus on identifying a particular rhetorical device (e.g., Metaphor (Rohanian et al. 2020; Stowe, Utama, and Gurevych 2022) and Exaggeration (Wright and Augenstein 2021)), ad understanding requires not only detecting possibly multiple rhetorical devices simultaneously, but also determining what the metaphor is.

[2]https://www.youtube.com/watch?v=GyY15Jkkg2A&ab_channel=HEADSHAUSADS



(1) (2) (3) (4)

(5) (6) (7) (8)

Figure 3: An example of video ads [3]. The selected frame tokens from the video are illustrated. The hidden message is *Buying insurance can help anyone and provide life security*.

Videos

Another form of ads is videos. Videos, which encompass texts, images, and audios, are a more complex form of ads. Videos often convey messages and ideas to the viewer by presenting a story. Figure 2 shows the frames in a video ad, which has no text. In this video, the child first buys two cans of Coke from the vending machine. Then he puts both cans under his feet to increase his height so that he can click the Pepsi button. Finally he leaves with the Pepsi contentedly and discards the Coke. This video conveys the hidden message *Pepsi is better than Coke* via *Comparison*. Typically, the *Storytelling* strategy is used by advertisers in video ads to make potential customers feel like they can relate to the story. One needs to watch the entire video instead of a clip to know the story behind the ad. For instance, if one only watches the first half of the video (a child is buying Coke), the hidden message would be *Coke is delicious*.

More sophisticated video ads incorporate text, narration, and background music. As shown in the ad in Figure 3, the male protagonist is eager to help others every day: he helps the elderly to push carts, he helps beggars, and he delivers fruits to the poor. At the end of the video, there are a few questions (written in the subtitles and narrated by the actor), such as "what do people who help others desire the most?" The video finally gives the answer ("Buy insurance to help the hero"), which suggests that it is an insurance ad. The superficial message of the video is that the male protagonist is constantly helping others, but the hidden message is *Buying insurance can help anyone and provide life security*. After watching the first half of the video, people can only see that the male protagonist helps others. They must combine the subtitles with the narration at the end of the video to understand the final purpose of the ad. By using sad music and relating concepts (e.g., "begging" represents "poverty", "sunsets" represent "difficulty"), the whole video creates a negative atmosphere, seeking to warn the viewers that if they do not purchase insurance, they will become vulnerable and need help from others, and implicitly prompt them to purchase insurance. Here, the attempt to create a negative atmosphere is related to the *Emotional Appeal* strategy, where advertisers persuade the audience by evoking an emotional response. Compared to image ads, video ads require a deeper comprehension from humans. Specifically, humans need to (1) understand not only texts and images, but also audios and emotions; (2) perceive the whole video instead of a clip; (3) identify the symbols used and their related concepts (e.g., "flowers" are related to "hopes", "sunsets" are related to "difficulty").

Challenges For AI systems to understand video ads, the following research questions need to be addressed: *(1) Extracting information from multiple modalities, including texts, images, audios, and emotions.* While there is a vast amount of work on extracting multimodal features, the majority of these efforts have only been able to extract the features from two or three modalities (Khare et al. 2021; Huang et al. 2020). In ad understanding, AI systems should be able to encode not only all modalities, including *emotions*, but also extract the correlations among modalities since some multimodal features need to be understood as a whole.

(2) Understanding the whole story, since it is possible to misinterpret a video by understanding only a portion of it. Analyzing contextual information from a video remains a challenging problem. Even though some videos are only a few minutes long, they may contain thousands of frame tokens. Currently available tools for image captioning (Fei 2022; Gao et al. 2022; Fei 2021) are unable to capture contextual information over a long sequence. Furthermore, they are unable to automatically select the most important frames to describe, often overlooking many important frames.

(3) Relating symbols to their concepts (e.g., "sunsets" represent "difficulty", which evokes a negative emotional response). It is generally not easy to recognize symbols that are not among the ones that are attention-capturing or eye-catching. For instance, the sunset in the fifth frame token from Figure 3 is hidden in the background. A further complication is that a symbol can be used to refer to different con-

[3] https://www.youtube.com/watch?v=uaWA2GbcnJU&ab_channel=thailifechannel

cepts. For instance, while "sunsets" represent "difficulty" in this example, this symbol may represent "relaxing rest after a long day's work" in another scenario. While the correlations between symbols and concepts can be extracted from a knowledge graph (KG) (Chen, Hu, and Sun 2022; Halliwell 2022), mapping symbols to nodes in KGs is not trivial.

3 Related Work

Existing work on computational advertising Work on computational advertising has focused on user targeting, identification of deceptive ads, and ad generation (Zhang et al. 2024; Zhao et al. 2024). These efforts primarily involve understanding surface-level information in ads such as the semantics of words and visual elements, and often neglect the persuasion strategies used and the arguments advertisers tried to make. Recently, some studies have examined persuasion strategies in memes (Abaskohi et al. 2024; Chikoti, Mehta, and Modi 2024). Understanding the persuasive arguments being made is critical for recognizing the advertisers' objectives. However, there is little research on understanding persuasive arguments in multimodal inputs, including ads.

Potentially useful techniques for ad analysis Next, we discuss techniques and tasks that are potentially relevant for analyzing ad content. Visual objects and morphemes in ads can be extracted using methods for object detection (Wang et al. 2021; Chen et al. 2021a; Zhou et al. 2021; Jeong et al. 2021; Joseph et al. 2021) and named entity identification (Shen et al. 2022; Chen et al. 2022; Das et al. 2022; Chen et al. 2021c; Wang and Henao 2021). Though these methods are useful for extracting information, they do not model the relationships among the different pieces of extracted information, which are essential for analyzing movements in ads. Multimodal sentiment analysis (Ling, Yu, and Xia 2022; Wu et al. 2022c; Yu et al. 2020) can help detect emotions directly expressed in words and images, but are relatively weak at detecting emotions that are only implicitly or subtly expressed. Existing natural language generation systems, such as image captioners (Fei 2022; Gao et al. 2022; Fei 2021) and machine translation systems (Li et al. 2022; Chiang et al. 2022; Hu et al. 2022), could be employed to describe the surface content of a single video frame or clip. Although an important step for our task, they remain inadequate as far as extracting the hidden message from the entire video is concerned. Methods for hand gesture detection (D'Eusanio et al. 2022; Silpani, Suematsu, and Yoshida 2022) can help identify the state changes of objects in ads. Tools for propaganda identification (Vijayaraghavan and Vosoughi 2022; Wang et al. 2020) could analyze persuasion strategies. Datasets, including Coco (Lin et al. 2014) and ImageNet (Deng et al. 2009), may provide data that can assist in extracting image objects. Research is needed to determine their effectiveness for ad understanding.

4 Approach

Given the challenges in ad understanding, how we can address this task? We propose to employ state-of-the-art machine learning techniques, especially deep learning, as their success in NLP and computer vision.

Corpus Construction

As there are no available datasets focusing on our task, the first step involves constructing an annotated dataset. We recommend that each ad in the dataset be annotated with four types of information: (1) the surface message(s); (2) the persuasion strategies; (3) the visual/textual elements in the ad through which the persuasion strategies are realized; and (4) the hidden message(s). The surface information (e.g., stories in videos) is essential for identifying persuasion strategies and generating hidden messages, as machines need to know what ads express explicitly. Consider the first video ad example. Knowing "Coke is stepped upon, while Pepsi is held" contributes to identifying the *Comparison* strategy. Persuasion strategies can be borrowed from existing inventories (Singla et al. 2022; Rocklage, Rucker, and Nordgren 2018; Hogan 2010). For each persuasion strategy, we annotate the textual elements (e.g., words, clauses, sentences), the visual elements (e.g., image objects), or the artifacts (e.g., video frames, or video clips) through which the strategy is realized. For instance, we should annotate the bounding box of the person Einstein in the "Einstein" example, as the person Einstein is related to the *Testimonial* strategy. In the first video example, we should annotate the clip of the video describing "The kid is holding Pepsi", as the clip is related to the *Comparison* strategy. Finally, for hidden messages, annotators should produce a textual description of each message implicitly conveyed the in ad. Note that some ads have multiple messages, while others have none.

The next question is: how can we resolve annotation conflicts? There are a thousand Hamlets in a thousand people's eyes. The same ad may be understood differently by different individuals. We think this inconsistency is acceptable. We argue that the annotated strategies or hidden messages may be different, but they would not be contradictory. Advertisers may employ ads to convey different messages to different groups of consumers, and AI systems should learn all these possibilities. Here, we suggest setting a threshold value to solve annotation conflicts. As mentioned before, it is not necessary for different annotators to annotate the same ad exactly the same. We could calculate inter-annotator agreement values and set different threshold values, then test the performance of models in the experiments with respect to different threshold values. After that, we could select a threshold value that does not cause significant changes in experimental performance.

Model Architecture

Given a dataset mentioned before, we can train a model to perform ad understanding in a supervised fashion. Nevertheless, there are several considerations that should be taken into account when designing the model architecture.

Unveiling hidden messages As mentioned in the introduction, one of the key challenges involves the subtask of identifying the hidden message(s) conveyed in an ad, if any. While it may be possible to train a model that directly maps an ad to its hidden message(s), it is conceivably a very challenging learning task. An alternative approach that we believe is promising is *argument mining*, which is a sub-

field of NLP where the goal is to identify the argumentative structure of a text document. Argument mining involves (1) identifying the argument components (i.e., the claims and premises/evidences) and linking them (i.e., determining which premises support which claim) to form an *argument tree*, where the root node corresponds to the hidden message, the children of the root node correspond to the claims/evidences that support the hidden messages, and each grandchild of the root node correspond to claims/evidences that support their parent node. Each leaf node of the tree corresponds to either a visual/text element that can be extracted from the ad or a piece of background knowledge relevant to ad interpretation.

While argument mining in NLP typically involves extracting argument components that are explicitly expressed in text, the key challenge involved in building argument trees for ads is that some claims in the tree (i.e., the non-leaf nodes) may not be explicitly expressed. To address this challenge, we propose to build the tree in a bottom-up fashion (e.g., Ng and Li (2023)). Specifically, we can first identify the surface messages (using an image captioner), extract text spans and visual elements from the ad, and obtain relevant background knowledge from external KGs or multimodal large language models (LLMs). These elements form the leaves of the argument trees. Then, we can use a reasoner (which can be an existing LLM or one that is fine-tuned for our task) to generate possible conclusions that can be derived from any subsets of these leaf nodes. Each of these conclusions will then serve as the parent of its supporting evidences (i.e., the leaf nodes). These conclusions can further be combined with each other or with any subset of the leaf nodes to derive further conclusions, allowing us to incrementally build an argument tree in a bottom-up fashion. This process can continue until we reach the root node, which, as mentioned above, corresponds to the hidden message.

To exemplify, consider building an argument tree for the ad in Figure 1(a). First, we identify the leaves, which correspond to the key text spans (such as "mentos") and visual elements (such as the object "Einstein") and extra knowledge (Einstein is an intelligent scientist). Next, from the text spans and the visual elements, we can infer that "Einstein is eating Mentos", which becomes the parent of these leaves. Combining this with the background knowledge "Einstein is an intelligent scientist", we can infer the hidden message "Eating Mentos will make you as smart as Einstein", which corresponds to the root node of the tree.

Building an argument trees for ads is not as straightforward as described. During tree construction, there are situations where the conclusion (i.e., parent node) derived from the children nodes corresponds to a metaphor, as in the ad in Figure 1(b), where we need to draw the conclusion that "a blurred woman" and "a clear woman" are metaphorically linked to "your phone" and "iPhone" respectively. Deriving conclusions that are metaphors is a largely unexplored task, especially from examples like this where the metaphor is derived by combining information from two modalities.

Encoding modalities As our input is multimodal, it is natural to train a multimodal model to encode all modalities. Off-the-shelf encoders like text encoders (e.g., SpanBERT), visual encoders (e.g., ResNet) and multimodal encoders (e.g., ViLBERT) could be employed (Wu et al. 2022a; Duan et al. 2024; Su et al. 2022; Zheng et al. 2021). As noted before, external relevant knowledge (e.g., background knowledge) may be needed to understand ads. If the extra knowledge is obtained from text or images, it can also be encoded using the aforementioned encoders. However, some external knowledge is not extracted from traditional modalities. For instance, information extracted from KGs corresponds to the graph modality and needs to be encoded using tailored encoders. Specifically, since graphs contain not only nodes but also relations, we need to adopt representation tools like TransE (Cai et al. 2018) that are tailored to KGs.

Inter- and intra-modal alignments are essential for comprehending multimodal information and can be investigated through existing multimodal alignment techniques. For example, the text span "tank" can have ambiguous meanings, as it could refer to either a tank top or a military tank. By aligning different modalities, the image modality can help disambiguate its meaning. However, current multimodal alignment methods can only identify obvious multimodal relationships, so it is worth further exploring how to align more *subtle* inter- and intra-modal connections, such as aligning "iphone" and "your phone" in Figure 1(b) with the clear person and the blurred person respectively.

Injecting knowledge Ad understanding may require a vast amount of background knowledge. One type of background knowledge is commonsense knowledge. While commonsense knowledge has been used extensively used in different NLP tasks in recent years, the kind of commonsense has largely be restricted to commonsense about entities (e.g., a celebrity's gender or hometown) and events (a bombing event is typically followed by a killing event) as well as the relationships between them. Unlike in these tasks, which rely primarily on what we refer to as *concrete* commonsense, in ad understanding it is not uncommon for us to additionally require *abstract commonsense*. Different types of abstract commonsense exist, such as those related to *color* and *plants*. For example, as shown in Figure 3, the rising sun symbolizes hope, whereas blooming flowers represent a pleasant mood. Understanding the emotions being conveyed is essential to properly understanding this and other ads.

Where can abstract commonsense be extracted from? There are already many knowledge bases for concrete commonsense, such as Wikipedia, but to our knowledge, few focuses on abstract commonsense. How to quickly assemble a large, high-quality knowledge base of abstract commonsense sense is a research question that needs to be addressed.

5 Conclusion

We introduced the task of ad understanding. A solution to this task could alleviate advertising manipulation by unveiling the hidden messages underlying an ad for its target audience. We believe that ad understanding can pose long-term challenges to AI researchers, and that the deep understanding needed by an ad understanding system could bring machine perception one step closer to human perception.

Acknowledgments

We thank the reviewers for their valuable comments on the earlier draft of this paper. This work was supported by National Natural Science Foundation of China (No. 62406139), State Key Laboratory for Novel Software Technology at Nanjing University (KFKT2023A07, KFKT2024A07, ZZKT2024B02).

References

Abaskohi, A.; Aghdam, A. D.; Wang, L.; and Carenini, G. 2024. BCAmirs at SemEval-2024 Task 4: Beyond Words: A Multimodal and Multilingual Exploration of Persuasion in Memes. In Ojha, A. K.; Dogruöz, A. S.; Madabushi, H. T.; Martino, G. D. S.; Rosenthal, S.; and Rosá, A., eds., *Proceedings of the 18th International Workshop on Semantic Evaluation, SemEval@NAACL 2024, Mexico City, Mexico, June 20-21, 2024*, 1412–1423. Association for Computational Linguistics.

Aberathne, I.; and Walgampaya, C. 2021. Real Time Mobile Ad Investigator: An Effective and Novel Approach for Mobile Click Fraud Detection. *Comput. Informatics*, 40(3).

Andrews, M.; Van Leeuwen, M.; Van Baaren, R.; and Plant, B. 2013. *Hidden persuasion: 33 psychological influence techniques in advertising*. Bis Publishers.

Blauth, T. F.; Gstrein, O. J.; and Zwitter, A. J. 2022. Artificial Intelligence Crime: An Overview of Malicious Use and Abuse of AI. *IEEE Access*, 10: 77110–77122.

Cai, R.; Liu, Y.; Zhang, M.; and Ma, S. 2018. Translating Embeddings for Modeling Query Reformulation. In Zhang, S.; Liu, T.; Li, X.; Guo, J.; and Li, C., eds., *Information Retrieval - 24th China Conference, CCIR 2018, Guilin, China, September 27-29, 2018, Proceedings*, volume 11168 of *Lecture Notes in Computer Science*, 3–15. Springer.

Camara, M. K.; Postal, A.; Maul, T. H.; and Paetzold, G. H. 2024. Can lies be faked? Comparing low-stakes and high-stakes deception video datasets from a Machine Learning perspective. *Expert Syst. Appl.*, 249: 123684.

Chacon, I. A.; and Sosnovsky, S. A. 2022. What's in an Index: Extracting Domain-specific Knowledge Graphs from Textbooks. In *WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022*, 966–976. ACM.

Chen, C.; Li, J.; Zheng, Z.; Huang, Y.; Ding, X.; and Yu, Y. 2021a. Dual Bipartite Graph Learning: A General Approach for Domain Adaptive Object Detection. In *2021 IEEE/CVF International Conference on Computer Vision, ICCV 2021, Montreal, QC, Canada, October 10-17, 2021*, 2683–2692. IEEE.

Chen, J.; Liu, Q.; Lin, H.; Han, X.; and Sun, L. 2022. Few-shot Named Entity Recognition with Self-describing Networks. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, 5711–5722.

Chen, L.; Yang, T.; Zhang, X.; Zhang, W.; and Sun, J. 2021b. Points As Queries: Weakly Semi-Supervised Object Detection by Points. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, 8823–8832. Computer Vision Foundation / IEEE.

Chen, S.; Aguilar, G.; Neves, L.; and Solorio, T. 2021c. Data Augmentation for Cross-Domain Named Entity Recognition. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, 5346–5356.

Chen, X.; Hu, Z.; and Sun, Y. 2022. Fuzzy Logic Based Logical Query Answering on Knowledge Graphs. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, 3939–3948. AAAI Press.

Chiang, T.; Chen, Y.; Yeh, Y.; and Neubig, G. 2022. Breaking Down Multilingual Machine Translation. In *Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland, May 22-27, 2022*, 2766–2780.

Chikoti, S.; Mehta, S.; and Modi, A. 2024. IITK at SemEval-2024 Task 4: Hierarchical Embeddings for Detection of Persuasion Techniques in Memes. In Ojha, A. K.; Dogruöz, A. S.; Madabushi, H. T.; Martino, G. D. S.; Rosenthal, S.; and Rosá, A., eds., *Proceedings of the 18th International Workshop on Semantic Evaluation, SemEval@NAACL 2024, Mexico City, Mexico, June 20-21, 2024*, 1779–1787. Association for Computational Linguistics.

Cialdini, R. B.; and Cialdini, R. B. 2007. *Influence: The psychology of persuasion*, volume 55. Collins New York.

Conia, S.; Lee, D.; Li, M.; Minhas, U. F.; and Li, Y. 2024. Enhancing Machine Translation Experiences with Multilingual Knowledge Graphs. In Wooldridge, M. J.; Dy, J. G.; and Natarajan, S., eds., *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024, Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence, IAAI 2024, Fourteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2014, February 20-27, 2024, Vancouver, Canada*, 23781–23783. AAAI Press.

Das, S. S. S.; Katiyar, A.; Passonneau, R. J.; and Zhang, R. 2022. CONTaiNER: Few-Shot Named Entity Recognition via Contrastive Learning. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, 6338–6353.

Deng, J.; Dong, W.; Socher, R.; Li, L.; Li, K.; and Fei-Fei, L. 2009. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), 20-25 June 2009, Miami, Florida, USA*, 248–255. IEEE Computer Society.

D'Eusanio, A.; Pini, S.; Borghi, G.; Simoni, A.; and Vezzani, R. 2022. Unsupervised Detection of Dynamic Hand Gestures from Leap Motion Data. In *Image Analysis and Processing - ICIAP 2022 - 21st International Conference,*

Lecce, Italy, May 23-27, 2022, Proceedings, Part I, volume 13231 of *Lecture Notes in Computer Science*, 414–424. Springer.

Duan, Z.; Wang, F.; Wang, B.; Luo, G.; and Jiang, Z. 2024. An Adapted ResNet-50 Architecture for Predicting Flow Fields of an Underwater Vehicle. *IEEE Access*, 12: 66398–66407.

Fei, Z. 2021. Memory-Augmented Image Captioning. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, 1317–1324. AAAI Press.

Fei, Z. 2022. Attention-Aligned Transformer for Image Captioning. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, 607–615. AAAI Press.

Feng, Y.; Wang, T.; Li, C.; Ng, V.; Ge, J.; Luo, B.; Hu, Y.; and Zhang, X. 2021. Don't Miss the Potential Customers! Retrieving Similar Ads to Improve User Targeting. In *Findings of the Association for Computational Linguistics: EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 16-20 November, 2021*, 1493–1503.

Gao, J.; Zhou, Y.; Yu, P. L. H.; Joty, S. R.; and Gu, J. 2022. UNISON: Unpaired Cross-Lingual Image Captioning. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, 10654–10662. AAAI Press.

Halliwell, N. 2022. Evaluating Explanations of Relational Graph Convolutional Network Link Predictions on Knowledge Graphs. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, 12880–12881. AAAI Press.

Hogan, K. 2010. *The psychology of persuasion: how to persuade others to your way of thinking*. Pelican Publishing.

Horák, A.; Sabol, R.; Herman, O.; and Baisa, V. 2024. Recognition of propaganda techniques in newspaper texts: Fusion of content and style analysis. *Expert Syst. Appl.*, 251: 124085.

Hu, J.; Hayashi, H.; Cho, K.; and Neubig, G. 2022. DEEP: DEnoising Entity Pre-training for Neural Machine Translation. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, 1753–1766.

Huang, G.; Pang, B.; Zhu, Z.; Rivera, C.; and Soricut, R. 2020. Multimodal Pretraining for Dense Video Captioning. In *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing, AACL/IJCNLP 2020, Suzhou, China, December 4-7, 2020*, 470–490.

Jeong, J.; Verma, V.; Hyun, M.; Kannala, J.; and Kwak, N. 2021. Interpolation-Based Semi-Supervised Learning for Object Detection. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, 11602–11611. Computer Vision Foundation / IEEE.

Joseph, K. J.; Khan, S. H.; Khan, F. S.; and Balasubramanian, V. N. 2021. Towards Open World Object Detection. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, 5830–5840. Computer Vision Foundation / IEEE.

Khare, Y.; Bagal, V.; Mathew, M.; Devi, A.; Priyakumar, U. D.; and Jawahar, C. V. 2021. MMBERT: Multimodal BERT Pretraining for Improved Medical VQA. *CoRR*, abs/2104.01394.

Li, B.; Lv, C.; Zhou, Z.; Zhou, T.; Xiao, T.; Ma, A.; and Zhu, J. 2022. On Vision Features in Multimodal Machine Translation. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, 6327–6337.

Lin, J.; Chen, B.; Wang, H.; Xi, Y.; Qu, Y.; Dai, X.; Zhang, K.; Tang, R.; Yu, Y.; and Zhang, W. 2024. ClickPrompt: CTR Models are Strong Prompt Generators for Adapting Language Models to CTR Prediction. In Chua, T.; Ngo, C.; Kumar, R.; Lauw, H. W.; and Lee, R. K., eds., *Proceedings of the ACM on Web Conference 2024, WWW 2024, Singapore, May 13-17, 2024*, 3319–3330. ACM.

Lin, T.; Maire, M.; Belongie, S. J.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; and Zitnick, C. L. 2014. Microsoft COCO: Common Objects in Context. In *Computer Vision - ECCV 2014 - 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V*, volume 8693 of *Lecture Notes in Computer Science*, 740–755. Springer.

Ling, Y.; Yu, J.; and Xia, R. 2022. Vision-Language Pre-Training for Multimodal Aspect-Based Sentiment Analysis. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, 2149–2159.

Maio, C. D.; Gallo, M.; Hao, F.; and Yang, E. 2021. Who and where: context-aware advertisement recommendation on Twitter. *Soft Comput.*, 25(1): 379–387.

Millar, K.; Cheng, A.; Chew, H. G.; and Lim, C. 2018. Deep Learning for Classifying Malicious Network Traffic. In *Trends and Applications in Knowledge Discovery and Data Mining - PAKDD 2018 Workshops, BDASC, BDM, ML4Cyber, PAISI, DaMEMO, Melbourne, VIC, Australia, June 3, 2018, Revised Selected Papers*, volume 11154 of *Lecture Notes in Computer Science*, 156–161. Springer.

Ng, V.; and Li, S. 2023. Multimodal Propaganda Processing. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence*, 15368–15375.

Quan, Y.; Ding, J.; Jin, D.; Yang, J.; Zhou, X.; and Li, Y. 2020. Representative Negative Instance Generation for Online Ad Targeting. In *CIKM '20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19-23, 2020*, 2177–2180. ACM.

Rasoolipour, S.; and Emamiifar, S. N. 2020. The effect of creative applicable methods on attracting the audience towards urban commercial advertisements from 2000 to 2020. *Int. J. Arts Technol.*, 12(3): 266–281.

Rocklage, M. D.; Rucker, D. D.; and Nordgren, L. F. 2018. Persuasion, emotion, and language: The intent to persuade transforms language via emotionality. *Psychological science*, 29(5): 749–760.

Rohanian, O.; Rei, M.; Taslimipoor, S.; and Ha, L. A. 2020. Verbal Multiword Expressions for Identification of Metaphor. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, 2890–2895.

Seyler, D.; Dembelova, T.; Corro, L. D.; Hoffart, J.; and Weikum, G. 2018. A Study of the Importance of External Knowledge in the Named Entity Recognition Task. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 2: Short Papers*, 241–246.

Shen, Y.; Wang, X.; Tan, Z.; Xu, G.; Xie, P.; Huang, F.; Lu, W.; and Zhuang, Y. 2022. Parallel Instance Query Network for Named Entity Recognition. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, 947–961.

Silpani, D. C.; Suematsu, K.; and Yoshida, K. 2022. A Feasibility Study on Hand Gesture Intention Interpretation Based on Gesture Detection and Speech Recognition. *J. Adv. Comput. Intell. Intell. Informatics*, 26(3): 375–381.

Singla, Y. K.; Jha, R.; Gupta, A.; Aggarwal, M.; Garg, A.; Bhardwaj, A.; Tushar; Krishnamurthy, B.; Shah, R. R.; and Chen, C. 2022. Persuasion Strategies in Advertisements: Dataset, Modeling, and Baselines. *CoRR*, abs/2208.09626.

Stowe, K.; Utama, P.; and Gurevych, I. 2022. IMPLI: Investigating NLI Models' Performance on Figurative Language. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, 5375–5388.

Su, H.; Shi, W.; Shen, X.; Xiao, Z.; Ji, T.; Fang, J.; and Zhou, J. 2022. RoCBert: Robust Chinese Bert with Multimodal Contrastive Pretraining. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, 921–931.

Sun, S.; Yu, L.; Zhang, X.; Xue, M.; Zhou, R.; Zhu, H.; Hao, S.; and Lin, X. 2021. Understanding and Detecting Mobile Ad Fraud Through the Lens of Invalid Traffic. In *CCS '21: 2021 ACM SIGSAC Conference on Computer and Communications Security, Virtual Event, Republic of Korea, November 15 - 19, 2021*, 287–303. ACM.

Thakare, Y. A.; and Walse, K. H. 2022. Automatic Caption Generation from Image: A Comprehensive Survey. In Jain, S.; Groppe, S.; and Mihindukulasooriya, N., eds., *Proceedings of the Workshop on Advances in Computational Intelligence, its Concepts & Applications (ACI 2022) co-located with International Semantic Intelligence Conference (ISIC 2022), Georgia Southern University, Savannah, United States, May 17-19, 2022*, volume 3283 of *CEUR Workshop Proceedings*, 282–293. CEUR-WS.org.

Thejas, G. S.; Boroojeni, K. G.; Chandna, K.; Bhatia, I.; Iyengar, S. S.; and Sunitha, N. R. 2019. Deep Learning-based Model to Fight Against Ad Click Fraud. In *Proceedings of the 2019 ACM Southeast Conference, ACM SE '19, Kennesaw, GA, USA, April 18-20, 2019*, 176–181. ACM.

Touahri, I.; and Mazroui, A. 2024. Survey of machine learning techniques for Arabic fake news detection. *Artif. Intell. Rev.*, 57(6): 157.

Vijayaraghavan, P.; and Vosoughi, S. 2022. TWEETSPIN: Fine-grained Propaganda Detection in Social Media Using Multi-View Representations. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL 2022, Seattle, WA, United States, July 10-15, 2022*, 3433–3448.

Wang, R.; and Henao, R. 2021. Unsupervised Paraphrasing Consistency Training for Low Resource Named Entity Recognition. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, 5303–5308.

Wang, R.; Tang, D.; Duan, N.; Zhong, W.; Wei, Z.; Huang, X.; Jiang, D.; and Zhou, M. 2020. Leveraging Declarative Knowledge in Text and First-Order Logic for Fine-Grained Propaganda Detection. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*, 3895–3903.

Wang, X.; Huang, T. E.; Liu, B.; Yu, F.; Wang, X.; Gonzalez, J. E.; and Darrell, T. 2021. Robust Object Detection via Instance-Level Temporal Cycle Confusion. In *2021 IEEE/CVF International Conference on Computer Vision, ICCV 2021, Montreal, QC, Canada, October 10-17, 2021*, 9123–9132. IEEE.

Wright, D.; and Augenstein, I. 2021. Semi-Supervised Exaggeration Detection of Health Science Press Releases. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, 10824–10836.

Wu, C.; Li, L.; Liu, Z.; and Zhang, X. 2022a. Machine Reading Comprehension Based on SpanBERT and Dynamic Convolutional Attention. In *Proceedings of the 4th International Conference on Advanced Information Science and System, AISS 2022, Sanya, China, November 25-27, 2022*, 41:1–41:5. ACM.

Wu, H.; Li, X.; Li, L.; and Wang, Q. 2022b. Propaganda Techniques Detection in Low-Resource Memes with Multi-

Modal Prompt Tuning. In *IEEE International Conference on Multimedia and Expo, ICME 2022, Taipei, Taiwan, July 18-22, 2022*, 1–6. IEEE.

Wu, Y.; Zhao, Y.; Yang, H.; Chen, S.; Qin, B.; Cao, X.; and Zhao, W. 2022c. Sentiment Word Aware Multimodal Refinement for Multimodal Sentiment Analysis with ASR Errors. In *Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland, May 22-27, 2022*, 1397–1406.

Yang, P.; Li, L.; Luo, F.; Liu, T.; and Sun, X. 2019. Enhancing Topic-to-Essay Generation with External Commonsense Knowledge. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, 2002–2012.

Yu, W.; Xu, H.; Meng, F.; Zhu, Y.; Ma, Y.; Wu, J.; Zou, J.; and Yang, K. 2020. CH-SIMS: A Chinese Multimodal Sentiment Analysis Dataset with Fine-grained Annotation of Modality. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, 3718–3727.

Zhang, H.; Ding, Z.; Dipu, M. S. K.; Lv, P.; Huang, Y.; Abdullahi, H. S.; Zhang, A.; Song, Z.; and Wang, Y. 2024. Identification of Illegal Outdoor Advertisements Based on CLIP Fine-Tuning and OCR Technology. *IEEE Access*, 12: 92976–92987.

Zhao, K.; Zhao, X.; Jin, Z.; Yang, Y.; Tao, W.; Han, C.; Li, S.; and Liu, L. 2024. Enhancing Baidu Multimodal Advertisement with Chinese Text-to-Image Generation via Bilingual Alignment and Caption Synthesis. In Yang, G. H.; Wang, H.; Han, S.; Hauff, C.; Zuccon, G.; and Zhang, Y., eds., *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2024, Washington DC, USA, July 14-18, 2024*, 2855–2859. ACM.

Zheng, R.; Chen, J.; Ma, M.; and Huang, L. 2021. Fused Acoustic and Text Encoding for Multimodal Bilingual Pretraining and Speech Translation. In *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, 12736–12746. PMLR.

Zhou, Q.; Yu, C.; Wang, Z.; Qian, Q.; and Li, H. 2021. Instant-Teaching: An End-to-End Semi-Supervised Object Detection Framework. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, 4081–4090. Computer Vision Foundation / IEEE.