

ORIGINAL ARTICLE

Knowledge-base of Semantic Relationships between Images based on FrameNet from Computational Semiotic Perspective

Masood Ghayoomi* 

Faculty of Linguistics, Institute for Humanities and Cultural Studies, Tehran, Iran.

Correspondence:

Masood Ghayoomi
Email: M.Ghayoomi@ihcs.ac.ir

Received: 08/Sep/2025
Revised: 07/Oct/2025
Accepted: 12/Oct/2025

How to cite:

Ghayoomi, M. (2026). Knowledge-base of Semantic Relationships between Images based on FrameNet from Computational Semiotic Perspective, *Journal of Sociolinguistics*, 9 (33), 37-65.
(DOI: [10.30473/il.2026.75736.1714](https://doi.org/10.30473/il.2026.75736.1714))

ABSTRACT

In the framework of semiotics, sign systems form the basis of human social and cultural activities for representing the knowledge hidden in signs. These signs may be manifested in the language or image system. Achieving abstract concepts in these two systems can help in classifying signs and representing them in the form of a knowledge base. In the present study, an attempt is made to use, within the framework of computational semiotics, the semantic relations of image signs that have been obtained in the framework of frame semantics obtained from linguistic signs of image captions to classify images. Then, this achievement is used to develop a knowledge base that contains these semantic relations between images. The results of this study indicate that extracting abstract concepts from frame-based semantic representation of image captions can help both the semantic classification of images and determine the semantic relations of images. The classification and relationship of images can be expressed as a triple containing the type of relationship and two image elements. This representation can be used in the construction of a knowledge base, as well as in the conceptual search of images. To conduct this research, the Flickr30k corpus, which contains an image and its relevant five captions written in English, is used.

KEYWORDS

Computational semiotics, frame semantics, image classification, knowledge base, semantic relationship of images, semantic elements.



Introduction

From the perspective of semiotics, all human social and cultural activities are based on a set of sign systems. These signs can manifest in the form of words, images, sounds, objects and behavior. Semiotics deals with how meanings are constructed and how reality is represented through signs, sign systems and processes of significant (Chandler, 2022: 2). In other words, semiotics deals with the representation of hidden knowledge in signs. In the structuralist approach, the identity of signs is formed in relation to other signs within a system. de Saussure (1983: 80) considers the value of signs to depend on the relationship of that sign to other signs and the position of that sign within that semiotic system. These signs can be words in the language system or elements in the system of an image.

Reading a sentence or seeing an image brings a meaning to mind. In the semiotic system of language, the meaning of a sentence is derived from the meanings of the words in that sentence in terms of paradigmatic and syntagmatic axes of each word. In the semiotic system of images, the concept of an image can also be obtained by composing the elements in the image and the semantic features of each element in that image. It is worth noting that just as in the language system, a sentence may have an implicit meaning in addition to its explicit meaning, this feature is also true in the image system; therefore, an image can have an explicit or implicit meaning.

The current research aims at categorizing images based on abstract concepts by using linguistic and image signs and extracting semantic elements from signs in the semiotic system. “Attention” to a part of an image sign and conveying the meaning by linguistic sign related to the attention is the key towards extracting the abstract concepts. The existence of this abstract information can lead to developing a knowledge base based on the semiotic concepts. Extracting these elements, whether explicit or implicit, and mapping them to images and text leads to the semantic understanding of images and text. It goes without saying that each element has its own unique semantic roles that deepen the semantic understanding of the text and image. These elements, which are obtained from the text and image, have the ability to be read as image and text in a common abstract space between the text and image. In this article, we will try to extract the abstract concepts and classify images based on the concepts.

Methodology

To perform this research to find the relationship between the language and image signs, we need to caption images by paying the “attention” to a part of an image and not the whole image. This means that it is possible to describe an image differently, using different vocabulary and syntactic constructions. Then, based on the captions, we find the main verb in the captions and extract the abstract concept of the verb from FrameNet which is constructed based on frame semantics to represent semantic properties of linguistic signs. Using the abstract semantic concept, we classify image signs and find the abstract semantic relations between the images.

To reach the goals, we use the Flickr30k dataset developed by Young et al. (2014). This dataset contains more than 30,000 images and relevant five captions which are written by human annotators and collected through crowd sourcing. The captions are unique; therefore, there is no repetition on the captions.

Findings

In this research, we extract the abstract concept of the main verbs of the captions from FrameNet and classify the images based on the abstract concepts. Since the frames vary frame elements in terms of type and number, we fix the type and number of frame elements for all frames. Moreover, we consider the frame name as the relationship element between two images, and define it as a triple logical formula.

Results and Conclusion

A) By paying “attention” to different parts of an image, we can explore the elements and

properties of each image. Therefore, the caption of each image is not necessarily one sentence .


- B) Each image caption has an abstract and general concept that can be used as a feature to find the relationship between images and classify them. In the present study, we achieved this goal and presented the result in the form of a knowledge-base, using triples .
- C) Finding an abstract concept and displaying it in the form of logical relationships paves the way to use this representation method in semantic search and inference in the knowledge-base .
- D) Sometimes abstract semantic relationships between images may be more than one concept. This method of analysis can be of great help in finding the semantic relationships between images.

Keywords

Computational semiotics, frame semantics, image classification, knowledge base, semantic relationship of images, semantic elements.

«مقاله پژوهشی»

پایگاه دانش روابط معنایی تصاویر مبتنی بر شبکه‌ی قالبی با رویکرد نشانه‌شناسی رایانشی

مسعود قیومی*  iD

۱دانشیار پژوهشگاه علوم انسانی و مطالعات فرهنگی، تهران، ایران

نویسنده مسئول:

مسعود قیومی

رایانامه: M.Ghayoomi@ihcs.ac.ir

تاریخ دریافت: ۱۴۰۴/۰۶/۱۷

تاریخ بازنگری: ۱۴۰۴/۰۷/۱۵

تاریخ پذیرش: ۱۴۰۴/۰۷/۲۰

استناد به این مقاله:

قیومی، مسعود (۱۴۰۵). پایگاه دانش روابط معنایی تصاویر مبتنی بر شبکه‌ی قالبی با رویکرد نشانه‌شناسی، فصلنامه زبان‌شناسی اجتماعی، ۹ (۳۳)، ۶۵-۳۷.

(DOI: [10.30473/il.2026.75736.1714](https://doi.org/10.30473/il.2026.75736.1714))

چکیده

در چارچوب نشانه‌شناسی، نظام‌های نشانه‌ای اساس فعالیت‌های اجتماعی و فرهنگی انسان را برای بازنمایی دانش نهفته در نشانه‌ها شکل می‌دهد. این نشانه‌ها ممکن است در نظام زبان یا تصویر تجلی یابد. دستیابی به مفاهیم انتزاعی در این دو نظام می‌تواند در دسته‌بندی نشانه‌ها و بازنمایی آن در قالب یک پایگاه دانش کمک نماید. در پژوهش حاضر تلاش می‌شود در چارچوب نشانه‌شناسی رایانشی، روابط معنایی نشانه‌های تصویری که در چارچوب معناشناسی قالبی از نشانه‌های زبانی حاصل از شرح‌های نوشته‌شده برای تصاویر به دست آمده‌است، برای دسته‌بندی تصاویر استفاده گردد. سپس، از این دستاورد، برای تهیه یک پایگاه دانش که حاوی این روابط معنایی میان تصاویر است استفاده گردد. نتایج حاصل از این پژوهش بیانگر این نکته است که استخراج مفاهیم انتزاعی حاصل از بازنمایی معنایی قالبی از شرح تصاویر می‌تواند ضمن کمک به دسته‌بندی معنایی تصاویر، ارتباطات معنایی تصاویر را مشخص کند. این دسته‌بندی و ارتباط تصاویر می‌تواند به صورت یک سه‌تایی که حاوی نوع رابطه و دو عنصر تصویر است بیان گردد و ضمن کاربرد در ساخت یک پایگاه دانش، در جستجوی مفهومی تصاویر کاربرد داشته باشد. برای انجام این پژوهش، از پیکره Flickr30k که حاوی تصویر و پنج شرح نگارش شده به زبان انگلیسی است استفاده شده است.

واژه‌های کلیدی

نشانه‌شناسی رایانشی، معناشناسی قالبی، دسته‌بندی تصویر، پایگاه دانش، رابطه معنایی تصاویر، عناصر معنایی.

مقدمه

ویژگی‌های معنایی هر یک از عناصر موجود در آن تصویر به دست آورد. شایان ذکر است همان‌گونه که در نظام زبان، یک جمله ممکن است علاوه بر معنای صریح، معنای ضمنی داشته باشد، این ویژگی در نظام تصویر نیز صادق است؛ به طوری که یک تصویر می‌تواند یک معنای صریح یا یک معنای ضمنی داشته باشد.

هدف از انجام این پژوهش، بازنمایی معنایی نشانه‌ها در نگاه به نشانه‌شناسی از دیدگاه علوم طبیعی و دسته‌بندی تصاویر براساس یک مفهوم انتزاعی با بهره‌گیری از نشانه‌های زبانی و تصویری و استخراج عناصر معنایی از نشانه‌ها در نظام نشانه‌شناسی است. «توجه»^۵ به بخشی از یک نشانه‌ی تصویری و انتقال معنا از طریق نشانه‌ی زبانی مرتبط با آن «توجه»، کلید استخراج مفاهیم انتزاعی است. وجود این اطلاعات انتزاعی می‌تواند به تهیه‌ی یک پایگاه دانش براساس مفاهیم نشانه‌شناسی منجر گردد. استخراج این عناصر، چه صریح و چه ضمنی^۶، و نگاشت آنها به تصویر و متن، به درک معنایی تصویر و متن می‌انجامد. ناگفته نماند که هر یک از عناصر خود دارای نقش‌های معنایی منحصر به فرد است که به درک معنایی متن و تصویر تعمیق می‌بخشد. این عناصر که از متن و تصویر به دست می‌آید، در یک فضای مشترک انتزاعی میان متن و تصویر قابلیت خوانش به تصویر و متن را دارد. در این مقاله می‌کوشیم که به استخراج مفاهیم انتزاعی و دسته‌بندی این تصاویر براساس این مفاهیم بپردازیم. از این‌رو، از پیکره‌ی Flickr30k که حاوی تصویر و شرح تصاویر به زبان انگلیسی است استفاده می‌کنیم.

ساختار این مقاله به این صورت است که در بخش 2، به پیشینه‌ی مطالعات انجام‌پذیرفته در حوزه‌ی نشانه‌شناسی متن و تصویر می‌پردازیم. در بخش 3، چارچوب نظری پژوهش مشخص می‌گردد. در ادامه‌ی این بخش، ارتباط میان دو نظام زبان و تصویر بحث و بررسی می‌شود و مفهوم «توجه» معرفی می‌گردد. برای یافتن مفهوم انتزاعی تصاویر از معناشناسی قالبی^۷ استفاده می‌کنیم. ویژگی‌های داده‌های پژوهش در بخش‌های 4 توضیح داده شده و در بخش 5 تحلیل می‌شود. در نهایت، در بخش 6، به جمع‌بندی و نتیجه‌گیری پژوهش می‌پردازیم.

از نگاه نشانه‌شناسی، تمامی فعالیت‌های اجتماعی و فرهنگی انسان براساس مجموعه‌ای از نظام‌های نشانه‌ای بنیان نهاده شده‌است. مونر^۱ (۲۰۲۲: ۸-۹) سه نگاه کلی را در مطالعات نشانه‌شناسی معرفی کرده‌است: الف: فلسفی، ب: زبان‌شناسی و ج: طبیعی. دو نگاه فلسفی و زبان‌شناسی برای اهل علوم انسانی شناخته‌شده‌تر است و نگاه به نشانه‌شناسی از دیدگاه علوم طبیعی کمتر مورد توجه قرار گرفته‌است. در نگاه طبیعی به نشانه‌شناسی، موجودات زنده، اشیاء و نمادها همچون زبان، شمایل و غیره، حاملان معنا در طبیعت است. از این‌رو، می‌توان نشانه‌شناسی را به عنوان دنباله‌رو یا الهام گرفته از، حداقل تا حدی، روش‌های علوم طبیعی که شامل مشاهده، آزمایش، صوری‌سازی است، معرفی نمود. به عبارتی دیگر، در نگاه به نشانه‌شناسی از دیدگاه علوم طبیعی، نشانه‌ها که می‌تواند به صورت واژه، تصویر، صوت، شیء و رفتار تجلی یابد با روش‌شناسی مطرح در حوزه‌ی علوم طبیعی قابل بررسی است. فناوری اطلاعات در حوزه‌ی علوم طبیعی قابل بررسی بنا بر این، می‌توان با امکانات فناوری اطلاعات به بررسی نشانه‌شناختی پرداخت که با اصطلاح «نشانه‌شناسی رایانشی»^۲ (مونر، ۲۰۲۲) مطرح می‌گردد.

نشانه‌شناسی به چگونگی ساخته‌شدن معانی و چگونگی بازنمایی واقعیت از طریق نشانه‌ها، نظام‌های نشانه‌ای و فرآیندهای دلالت می‌پردازد (چندلر^۳، 2022: 2). به عبارتی دیگر، نشانه‌شناسی به بازنمایی دانش نهفته در نشانه‌ها می‌پردازد. در رویکرد ساختگرای، هویت نشانه‌ها در ارتباط با دیگر نشانه‌ها در درون یک نظام شکل گرفته‌است. سوسور^۴ (۱۹۸۳: ۸۰) ارزش نشانه‌ها را وابسته به رابطه‌ی آن نشانه با سایر نشانه‌ها و جایگاه آن نشانه در درون آن نظام نشانه‌ای می‌داند. این نشانه‌ها می‌تواند واژه‌ها در نظام زبان یا عناصر موجود در نظام یک تصویر باشد.

خواندن یک جمله یا دیدن یک تصویر، معنایی را به ذهن متبادر می‌سازد. در نظام نشانه‌ای زبان، معنای یک جمله از معانی واژه‌های آن جمله در محور جانشینی و همنشینی و نقش‌های معنایی هر یک از واژه‌ها حاصل می‌گردد. در نظام نشانه‌ای تصویر نیز می‌توان مفهوم یک تصویر را از کنار هم قرار دادن عناصر موجود در تصویر و

5. attention

۶. برای نمایش عناصر معنایی ضمنی در تصاویر از زیرخط استفاده می‌نماییم.

7. Frame semantics

1. Meunier, J.G.

2. computational semiotics

3. Chandler, D.

4. de Saussure, F.

پیشینه مطالعاتی

در حوزه نشانه‌شناسی تصویری و زبان، مطالعات متنوع و مستقلی از یکدیگر انجام شده‌است. یک رویکرد این است که به بررسی خود تصویر پرداخته شود و ویژگی‌های معنایی مانند شباهت و همسانی عناصر موجود در تصویر، فاصله و دوری و نزدیکی عناصر، کادربندی در تصویر، چپش عناصر در تصویر، زاویه نور و مانند آن که کپس^۱ (۱۴۰۳) به‌عنوان زبان تصویر معرفی کرده‌است از تصویر استخراج شده، به ارتباط بصری این عناصر پرداخته شود. رویکرد دیگر تحلیل گفتمان چندوجهی^۲ است که بازنمایی، ارتباط و تعامل را چیزی بیش از زبان می‌داند (جوویت^۳، ۲۰۰۹: ۱). حکیم و همکاران (۱۴۰۰) هنر را رابط ناگسستگی زبان و تصویر می‌داند. در تحلیل گفتمان چندوجهی، انسجام درونی متن و تصویر در یک بافت خاص (کرس^۴ و همکاران، ۱۴۰۱) از جنبه نشانه‌شناسی اجتماعی مورد مطالعه قرار می‌گیرد که در ادامه به بررسی بعضی از پژوهش‌های مرتبط در این حوزه می‌پردازیم. هاتفی و شعیری (۱۳۹۰) با رویکرد تحلیل گفتمان پدیدارشناختی، تعامل متن و تصویر در کتاب مصور «مردم معمولی» را مطالعه کرده، بینامتنیت میان متن و تصویر را مورد توجه قرار داده‌اند. رویکردهای دیگر نشانه‌شناسی تصویر و زبان می‌تواند در چارچوب یک نظریه مشخص یا کاربرد تحقق پذیرد.

پژوهش‌های نظری

صادقی (۱۳۹۱) در چارچوب نظریه ادغام مفهومی، به بررسی رابطه تصویر و نوشتار و درک آن پرداخته‌است. شکل‌گیری این رابطه و بازنمایی عملکرد آن همگی در ذهن مخاطب است. در این پژوهش از روش فوکونیه^۵ و ترنر^۶ (۲۰۰۲) برای نمایش عملکرد استعاری ذهن استفاده شده‌است. براساس مطالعات انجام‌شده، تصویر سبب تغییر معنای استعاری نوشتار می‌شود و تبدیل نوشتار به تصویر برای معناسازی به‌کار می‌رود.

مهدی‌زاده (۱۳۹۲) به مطالعه کنش عکاسی و خوانش عکس در چارچوب نشانه‌شناسی پرداخته‌است. وی معتقد است عکس نوعی متن نشانه‌ای است که با واقعیت و عناصر عالم بیرون از عکس در ارتباط است و معنای مستقلی را به

تأویل و تفسیر وا می‌گذارد. وی در بررسی خود به معانی صریح و ضمنی در عکاسی اشاره کرده‌است. عمادالدین (۱۳۹۶) از روش نشانه‌شناسی بارت^۷ (1968) برای بررسی تطابق متن خبر و عکس ۳۰ نمونه در سه شبکه خبری، العالم و پرس‌تی‌وی استفاده کرده‌است. براساس نتایج، ۱۶ مورد تطابق کامل، ۸ مورد بدون تطابق و ۷ مورد تا حدودی دارای تطابق بود.

دیانت (۱۳۹۶) در پژوهش خود در چارچوب نظری پیرس^۸ (۱۹۳۱)، به بررسی اهمیت به‌کارگیری نشانه‌شناسی در عکس‌های مفهومی و دلالت‌های معنایی در تصاویر پرداخته‌است. براساس بررسی نمونه‌های این پژوهش، این نتیجه به‌دست آمد که تصاویر از دیدگاه پیرس نشانه‌های شمایی به حساب می‌آید که نقش نماد یا نمایه پیدا کرده و در انتقال معنا و مفهوم اثر هنری نقش کلیدی را بازی می‌نماید.

عمادالدین (۱۳۹۹) به نشانه‌شناسی عکس خبری و رمزگشایی معنایی از آن پرداخته‌است. در پژوهش خود از دیدگاه نشانه‌شناسان پیرس (۱۹۳۱) و بارت (۱۹۶۸) به مطالعه عکس‌های صفحه اول سه روزنامه شرق، اعتماد و ایران که در یک روز منتشر شده‌است پرداخته‌است.

تنهایی اهری و همکاران (۱۳۹۹) در پژوهش خود از روش نشانه‌معناشناسی گفتمانی و نشانه‌شناسی اجتماعی کرس و ون‌لیوون^۹ (2021) برای تحلیل کتاب‌های درسی زبان انگلیسی Vision بهره گرفته‌اند که این چارچوب نظری متأثر از آثار و نگرش هلیدی^{۱۰} (۱۹۹۴) در تعریف سه نقش اصلی زبانی بوده‌است. در این پژوهش، تصاویر کتاب براساس معنای بازنمودی، تعاملی و ترکیبی بررسی شده‌است.

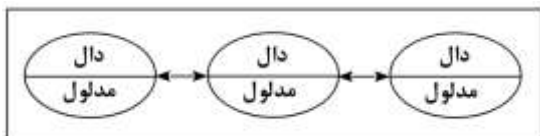
پژوهش‌های کاربردی

مطالعات انجام‌گرفته در حوزه متن و تصویر که نمونه‌هایی از آن در سطرهای پیشین بررسی شد، بیشتر به تبیین یک مسئله در چارچوب یک نظریه پرداخته‌است. دسته‌بندی تصاویر موضوع دیگری است که جزء اهداف پژوهش حاضر است. دسته‌بندی تصاویر بیشتر جنبه کاربردی دارد و با کمک فناوری اطلاعات انجام پذیرد. از این‌رو، استفاده از روش‌های یادگیری ماشین و استفاده از هوش مصنوعی در دسته‌بندی

7. Barthes, R.
8. Peirce, C.S.
9. Van Leeuwen, T.
10. Halliday, M. A. K.

1. Kepes, G.
2. Multi-modal discourse analysis
3. Jewitt, C.
4. Kress, G.
5. Fauconnier, G.
6. Turner, M.

نشانه با دیگر نشانه‌ها و جایگاه آن نشانه در درون آن نظام و نه نشانه‌های منفرد وابسته است (سوسور، 1983: 80). شایان ذکر است هر یک از نشانه‌ها دارای ارزشی است که از رابطه‌ی میان یک نشانه با دیگر نشانه‌ها در درون یک نظام تعیین می‌گردد (سوسور، 1983: 112-113). در شکل 1، رابطه‌ی میان نشانه‌ها در یک نظام به تصویر کشیده شده‌است (سوسور، 1983: 118).



شکل 1. رابطه‌ی نشانه‌ها در یک نظام

پیرس یک الگوی سه‌وجهی برای نشانه معرفی کرده‌است که عبارت است از بازنمون²، تفسیر³ و موضوع⁴. اگر بخواهیم مقایسه‌ای میان دیدگاه سوسور و پیرس به‌دست آوریم، بازنمون، همان دال است؛ تفسیر، مدلول است و موضوع، وجهی از نشانه‌شناسی پیرسی است که سوسور به آن نپرداخته‌است (سیلورمن⁵، 1983: 15). برای مثال، چراغ راهنمایی بازنمون است که موضوع آن توقف خورو بوده و تفسیر آن این است که خودرو باید با چراغ قرمز متوقف شود.

ارتباط میان نظام زبان و نظام تصویر

در بخش پیشین به آرای سوسور مربوط به نشانه در مکتب ساختگرایی پرداخته شد. براساس نظر سوسور، هر نشانه متشکل از دال و مدلول است و در یک نظام مشخص میان دال و مدلول ارتباط وجود دارد. با این نگاه نیز می‌توان در نظام تصویر به تحلیل نشانه‌های تصویر پرداخت؛ به این صورت که هر عنصر موجود در تصویر، یک نشانه است که متشکل از دال (تصویر گرافیکی) و مدلول (مفهومی که دال به آن دلالت می‌کند) است. به عبارتی دیگر، دیدن «گل» در یک تصویر به‌عنوان قسمتی از یک نشانه، حاوی مفهومی است که آن تصویر به آن دلالت می‌کند. بنابراین، در چارچوب نشانه‌شناسی، هریک از عناصر موجود در یک تصویر گرافیکی به‌عنوان یک نشانه در نظام تصویر، برابر است با هر یک از واژه‌ها به‌عنوان نشانه در نظام زبان.

تصاویر اجتناب‌ناپذیر است. این رویکرد سبب می‌شود پا از دسته‌بندی تصاویر معمولی فراتر رود و جنبه‌ی کاملاً کاربردی پیدا کند، مانند طبقه‌بندی معنایی تصاویر پزشکی براساس محتوای شکلی و بافتی که توسط پورقاسم و قاسمیان (۱۳۸۶) انجام شده‌است.

بخش‌بندی معنایی پیکسل‌های تصاویر با استفاده‌ی همزمان از اطلاعات سطح شی و سطح متن در تعیین برجسب‌های معنایی تصویر، پژوهش دیگری است که توسط رزاقی (۱۳۹۷) انجام پذیرفته‌است.

بازیابی معنایی تصویر نمونه‌ی دیگری از کاربرد پردازش الگوریتمی تصویر در فرایند جستجوی معنایی است که توسط باقری و همکاران (۱۳۹۹) انجام شده‌است.

پژوهش ژی¹ و همکاران (2020) منجر به تهیه‌ی پایگاه دانش تصویر و بازنمایی رابطه‌ی میان تصاویر شده‌است. پایگاه دانش می‌تواند در بازیابی اطلاعات و شیوه‌ی استدلال سامانه‌ی پرسش و پاسخ تصویری کاربرد داشته باشد که در این زمینه بارزی و کردجمشیدی (۲۰۲۴) پژوهش خود را انجام داده‌اند.

بلکاولو و همکاران (۲۰۲۴) در پژوهش خود تلاش کرده‌اند یک مجموعه‌ی داده ویدئویی چندوجهی را تهیه کنند که متشکل از تعدادی برنامه‌ی تلویزیونی مربوط به موضوع مسافرت در برزیل همراه با زیرنویس مصاحبه‌ی مجری به زبان انگلیسی است. سپس آنان با استفاده از بازنمایی معنایی شبکه‌ی قالبی، ارتباط میان متن و تصویر در هر فریم فیلم ویدئویی را به‌دست آورده و از این اطلاعات برای دسته‌بندی استفاده نموده‌اند.

چارچوب نظری

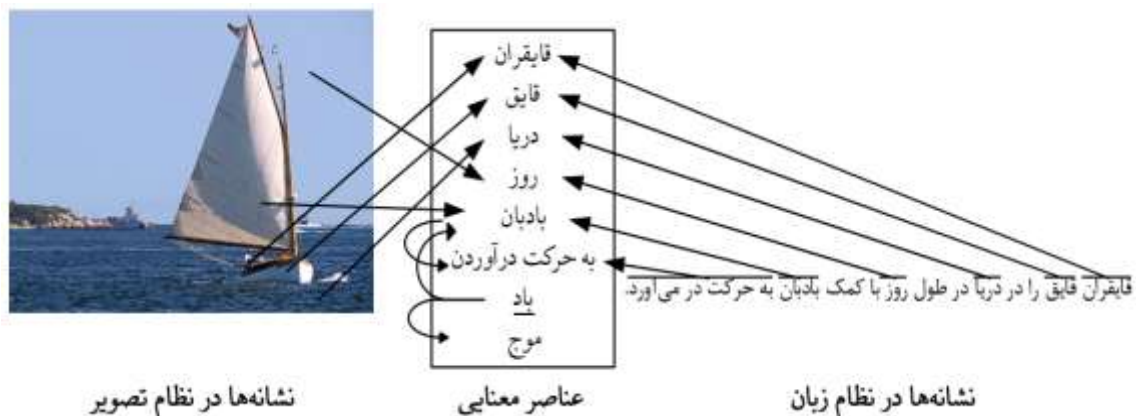
فردیناندو سوسور، زبان‌شناسی سویسی، و چارلز ساندرس پیرس، فیلسوف آمریکایی، که در یک دوره‌ی تاریخی ولی در دو قاره زیسته‌اند بنیان‌گذاران اصلی نشانه‌شناسی بوده‌اند. سوسور (1983: 67) نشانه را متشکل از دال (تصویر صوتی) و مدلول (مفهومی که دال به آن دلالت می‌کند) در یک نظام نشانه‌ای می‌داند. نکته‌ی قابل توجه در دیدگاه سوسوری به نظام و نشانه این است که از نظام شروع می‌کند تا به موضوع نشانه بپردازد. بنابراین، نظام در اولویت قرار دارد که از روابط بین نشانه‌ها حاصل شده‌است (سجودی، 1387: 24). سوسور (1983: 121) معتقد است در نظام زبان همه چیز نشانه‌ها وابسته به روابط است؛ و ارزش این نشانه‌ها به رابطه‌ی آن

2. Representation
3. Interpretation
4. Object
5. Silverman, K.

1. Xie, R.

آن با تصویر در چارچوب نشانه‌شناسی پرداخت. برای نمونه، در شکل ۲، یک جمله در مورد یک تصویر بیان شده است.

اگر برابری نظام زبان و تصویر که در بالا توضیح داده شد را بپذیریم، می‌توان به بررسی مفهوم یک جمله و ارتباط



شکل ۲. خوانش عناصر معنایی به متن و تصویر

را در دریا در طول روز با کمک بادبان به حرکت در می‌آورد.» بیان شده است. همان‌طور که در شکل ۲ ملاحظه می‌گردد مفهوم انتزاعی که در این تصویر نهفته است «عاملی برای تغییر موقعیت» است. برای این تغییر موقعیت، عناصر معنایی صریحی وجود دارد که میان نظام‌های زبان و تصویر مشترک است، مانند «قایقران»، «قایق»، «دریا»، «روز»، «بادبان» و «موج». همچنین عناصری معنایی دیگری در این مثال وجود دارد که در نظام زبان صریح و در نظام تصویر ضمنی است، مانند به‌حرکت‌درآوردن. اگرچه عنصر معنایی «به‌حرکت‌درآوردن» در جمله الصاق شده به تصویر، در نظام زبان به صراحت ذکر شده است، این عنصر به‌طور ضمنی در نظام تصویر قابل فهم است که از طریق «بادبان» افراشته و حرکت «موج» فهم می‌گردد. شایان ذکر است، بعضی از عناصر معنایی صریح ممکن است در جمله ذکر نشود ولی در نظام تصویر به‌طور ضمنی وجود داشته باشد، مانند «باد». عامل ایجاد «موج» و «به‌حرکت‌درآوردن» در شکل ۲، نشانه‌ای به‌نام «باد» است که وجود این «باد» در «بادبان» سبب «به‌حرکت‌درآوردن» این «قایق» شده است. به عبارتی دیگر، رابطه سببیت میان نشانه‌های «باد»، «بادبان» و «موج» که سبب «به‌حرکت‌درآوردن» این «قایق» شده است به‌طور ضمنی وجود دارد؛ درحالی‌که، فاعل دستوری این جمله که «قایق» را به حرکت در آورده است «قایقران» است.

واژه‌های این جمله که همان نشانه‌ها در نظام زبان است با مصادیق عینی موجود در تصویر که همان نشانه‌ها در نظام تصویر است مرتبط است. بنابراین، در فضایی میان متن و تصویر، مفاهیم انتزاعی و عناصر معنایی وجود خواهد داشت که بین این دو نظام، ارتباط برقرار می‌کند. یکی از ویژگی‌های این فضای معنایی این است که فقط حاوی واژه‌های محتوایی است و واژه‌های دستوری در این فضا جایی ندارد. ویژگی دیگر این فضا این است که این مفهوم انتزاعی حاوی عناصر معنایی صریح و ضمنی است؛ درحالی‌که ممکن است در نظام زبان یا تصویر نشانه‌ها به‌صورت صریح یا ضمنی موجود باشد.

همان‌طور که سوسور ارزش نشانه‌ها را در ارتباط یک نشانه با سایر نشانه‌ها می‌داند، این پدیده در نظام تصویر نیز دیده می‌شود به این مفهوم که عناصر موجود در یک تصویر که هر کدام به‌عنوان یک نشانه معرفی می‌شود، با دیگر عناصر موجود در تصویر در ارتباط است تا مفهوم تصویر در ذهن درک گردد. به‌عبارتی دیگر، هر یک از نشانه‌ها در یک بافتار قابل تفسیر است تا معانی صریح یا ضمنی آن نشانه مشخص گردد.

مفهوم این موضوع در مثال شکل ۲ توضیح داده می‌شود که شامل یک تصویر و توصیف زبانی آن تصویر است. این تصویر، حاوی یک قایق بادی با تعدادی افراد در داخل قایق است که در دریا در حال حرکت است. این عناصر در نظام تصویر در نظام زبان با جمله «قایقران قایق

توجه

جمله‌ی ذکرشده در شکل 2 فقط یکی از جملاتی است که می‌تواند برای توصیف تصویر مطرح گردد. در این تصویر، «ساحل صخره‌ای» و همچنین دو «کشتی کروز» که یکی در ساحل پهلو گرفته و دیگری که در وسط دریا در حال حرکت است نیز دیده می‌شود. هر توصیف دیگری به جز توصیف ذکرشده در این مثال، بیانگر تفاوت دیدگاه در درک تصویر است. به عبارتی دیگر، در نظام زبان و تصویر، در هر لحظه، گوینده یا بیننده دست به انتخاب تعدادی از نشانه‌های آن نظام می‌زند و در همان حد آن را درک می‌نماید که این فرایند «انتخاب توجه»^۱ نامیده می‌شود. «توجه» اصطلاحی است که توسط برادبنت^۲ (1953) و استیون^۳ (1953) در حوزه‌ی علوم اعصاب و روان‌شناختی شناختی معرفی شده‌است. به عبارتی دیگر، درک تصویر در نظام نشانه‌ای تصویر، توجه‌محور است و اینگونه نیست که تمام تصویر در هر لحظه به صورت کلی درک شود؛ بلکه در هر لحظه قسمتی از تصویر مورد توجه قرار گرفته و درک می‌گردد. در نتیجه، باتوجه به وجود ویژگی «توجه» ممکن است برای یک تصویر در نظام تصویر بیش از یک توصیف درک گردد و مابه‌ازای هر توصیف، یک جمله در نظام زبان ذکر شود.

شبکه‌ی قالبی

تحلیل ساختار معنایی یکی از روش‌های درک معنایی جمله است که کنش‌های انتزاعی را به همراه شرکت‌کنندگان کنش‌ها توصیف می‌کند. یکی از این روش‌ها، معناشناسی قالبی است که توسط چارلز فیلمور^۴ در دهه‌ی ۱۹۷۰ برای توصیف و بازنمایی معنایی در سطح انتزاعی معرفی شده‌است. یکی از خاستگاه‌های این چارچوب معناشناسی، دستور زبان حالت فیلمور^۵ (۱۹۶۸) است که با عنوان «دستور حالت» شناخته می‌شود. از نظر فیلمور، یک چارچوب به‌عنوان یک صحنه‌ی انتزاعی (یک صحنه‌ی طرح‌واره‌ای) در نظر گرفته می‌شود که دارای تعدادی شرکت‌کننده به‌عنوان موضوعات این گزاره و برخی جملات برای توصیف صحنه است. درواقع این چارچوب‌ها، ساختارهای مفهومی برای دانش پس‌زمینه

صحنه‌های انتزاعی است که توسط واحدهای واژگانی نمایش داده می‌شود و زمینه را برای عناصر کنش فراهم می‌کند. در نتیجه، برای «درک ساختار معنایی گزاره [مثلاً فعل] لازم است ویژگی‌های چنین صحنه‌ی طرح‌واره‌ای را درک کنیم» (فیلمور، ۱۹۸۲: ۱۱۵)؛ به طوری که بتوانیم نقش‌های معنایی هر یک از شرکت‌کنندگان صحنه را باتوجه به توصیف گزاره درک کنیم. فیلمور (۱۹۸۵) مفهوم «قالب» را برای «درک معنایی» در نظر گرفت تا مشخص کند که یک جمله برای شنونده در چه موقعیتی مناسب است تا تفسیر یک جمله را ارائه دهد.

در دیدگاه اخیر فیلمور، درک معنایی، ترکیبی است؛ به طوری که دانش واژه‌ها، عبارات و ساختارهای دستوری برای تفسیر در نظر گرفته می‌شود. بر اساس آنچه گفته شد، هر قالب، دو سطح دارد: یکی سطح مفهومی که در آن قالب، یک موقعیت طرح‌واره‌ای خاص را مدل‌سازی می‌کند و دانش پیش‌زمینه‌ای مرتبط با آن را نشان می‌دهد؛ و دیگری سطح زبانی است که در آن یک قالب به‌عنوان یک طبقه‌ی معنایی در نظر گرفته می‌شود و شامل تمام گزاره‌هایی می‌شود که قادر به بیان موقعیت هدف است. (پادو، ۲۰۰۷). در سطح مفهومی، نقش‌های انتزاعی به نام «نقش‌های معنایی» به شرکت‌کنندگان عمل در گزاره اختصاص داده می‌شود. چنین اطلاعاتی با کارکردهای نحوی موضوعات (مثلاً فاعل، مفعول) متفاوت است. فیلمور (۱۹۶۸؛ ۱۹۷۱) مجموعه‌ای از نقش‌های معنایی، مانند کنش‌گر^۶، تجربه‌گر^۷، ابزار^۸، غیرعامل^۹، واقعی^{۱۰}، منبع^{۱۱}، هدف^{۱۲}، مکان^{۱۳}، زمان^{۱۴}، مسیر^{۱۵} و هدف^{۱۶} را معرفی کرده‌است. این نقش‌های معنایی درواقع با «عناصر قالب»^{۱۸} در یک قالب مرتبط است که توسط برخی گزاره‌های خاص به نام «عناصر فراخوانی قالب»^{۱۹} به کار می‌رود.

در میان کاربردهای معناشناسی قالبی، یکی از مهم‌ترین

6. Padó, S.
7. agentive
8. experiencer
9. instrumental
10. dative
11. factitive
12. source
13. goal
14. locative
15. time
16. path
17. objective
18. Frame Elements (FE)
19. Frame Evoking Element (FEE)

1. selective attention
2. Broadbent, D. E.
3. Stevens, S. S.
4. Charles Fillmore
5. Fillmore, C.J.

جدول ۱. نمونه‌ای از ساختار اطلاعات مربوط به قالب COMMITMENT (پیمان) در شبکه قالبی

Frame: COMMITMENT	
تعریف	A Speaker makes a commitment to an Addressee to carry out some future action. This may be an action desirable (as with promise) or not desirable (as with threaten) to the Addressee.
عناصر قالب	<p>Speaker The Speaker is the person who commits him/herself to do something</p> <p>Addressee The Speaker's commitment can be made to an Addressee.</p> <p>Message An expression of the commitment made by the Speaker</p> <p>Topic The topic about which the Speaker makes a promise.</p> <p>Medium The Medium is the physical entity or channel used to transmit the Message.</p>
عناصر فراخوانی قالب	consent.v, covenant.n, covenant.v, oath.n, vow.n, pledge.n, pledge.v, promise.n, promise.v, swear.v, threaten.n, threaten.v, undertake.v
نشانه‌گذاری	<p>[Speaker Democratic audiences] had to consent [Message to this approach].</p> <p>[Speaker The politicians] made vague promises [Topic about independence].</p> <p>[Speaker he] [Message 'I'll be back , ' threatened.</p>

روش اجرا و داده‌های پژوهش

در بخش پیشین به چارچوب نظری پژوهش در نظام نشانه‌ای سوسوری برای نظام زبان و تصویر پرداخته شد و در ادامه نیز ویژگی «توجه» معرفی شد. همان‌طور که در مقدمه مقاله به آن اشاره شد، در پژوهش حاضر تلاش می‌شود نام قالب‌ها به‌عنوان مفهوم انتزاعی که از شرح ارائه‌شده (نشانه زبانی) برای یک تصویر (نشانه تصویری) که براساس ویژگی «توجه» به قسمتی از تصویر به‌دست آمده‌است از دیدگاه نشانه‌شناسی استخراج شود و از این اطلاعات برای دسته‌بندی تصاویر و تهیه پایگاه دانش استفاده گردد. منظور از قالب معنایی، همان فضای میانی است که در بین نظام زبان و نظام تصویر براساس یک مفهوم انتزاعی و عناصر معنایی متعلق به این مفهوم شکل می‌گیرد و حاوی واژه‌های

کاربردهای آن در حوزه فرهنگ‌نویسی و الزام به وجود یک طرح‌واره طبقه‌بندی خاص بر واژگان است (فیلمور، ۱۹۹۴). طرح فریم‌نت بر کلی^۱ (بیکر^۲ و همکاران، ۱۹۹۸) یک واژه‌نامه معنایی برای زبان انگلیسی است که در چارچوب معناشناسی قالبی گردآوری شده‌است. تهیه این واژه‌نامه کاربرد دیگر آن است. فریم‌نت بر چهار دسته واژه، یعنی افعال، اسم‌ها، صفات و حروف اضافه، تمرکز دارد. علاوه بر این، عبارات چندواژه‌ای، واژه‌های چند جزئی با فاصله و عبارات اصطلاحی نیز به‌عنوان یک واحد واژگانی تعریف شده‌است که در آن قالب‌های مناسب، بدون تجزیه و تحلیل ساختارهای داخلی به آنها اختصاص داده شده‌است (راپنهوفر^۳ و همکاران، ۲۰۰۶). هر قالب شامل ۵ نوع اطلاعات است: ارائه نام قالب به عنوان طبقه معنایی؛ تعریف قالب به زبان طبیعی؛ فهرست عناصر شرکت‌کننده در موقعیت‌ها و پیش‌فرض‌های موقعیت قالب معنایی؛ ارائه عناصر فراخوانی قالب که با الگوی بن‌واژه-نقطه-مقوله دستوری معرفی می‌شود؛ و ارائه چند مثال شاهد برچسب‌گذاری شده. در جدول ۱ می‌توان اطلاعات مربوط به قالب COMMITMENT (پیمان) را مشاهده نمود که در آن بر اساس تعریف داده‌شده، عناصر قالب فهرست شده و عناصر فراخوانی که ویژگی‌های یکسانی دارد در کنار مثال‌های نشانه‌گذاری شده ارائه شده‌است.

این نوع نشانه‌گذاری داده‌ها برای شناسایی نقش‌های معنایی در کاربردهای، مانند استخراج اطلاعات، ترجمه ماشینی آماری، خلاصه‌سازی خودکار متن و داده‌کاوی متن، بسیار مهم است (هرست^۴، ۱۹۹۹). علاوه بر این، چنین نشانه‌گذاری می‌تواند در سامانه‌های پرسش و پاسخ، سامانه‌های گفتگوی معنایی، ابهام‌زدایی معنایی واژه‌ها و مدل‌های زبانی احتمالاتی در سامانه‌های تشخیص گفتار (گیل‌دیا^۵ و جورافسکی^۶، ۲۰۰۲) به‌منظور داشتن مدل‌های زبانی بهتر استفاده شود. همان‌طور که در مثال جدول ۱ نشان داده شد، برای فهم یک جمله، علاوه بر در نظر گرفتن واحدهای واژگانی و روابط نحوی آنها، شناسایی قالب تخصیص داده‌شده به گزاره و همچنین تشخیص روابط مفهومی بین عناصر جمله با توجه به این قالب، مهم است.

1. Berkeley FrameNet
2. Baker, C.F.
3. Ruppenhofer, J.
4. Hearst, M.
5. Gildea, D.
6. Jurafsky, D.

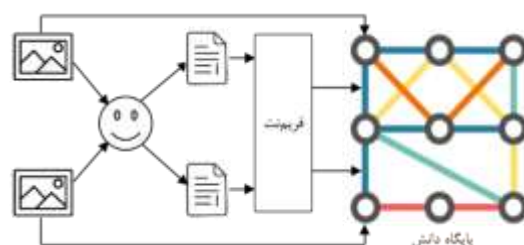
موجود در تصویر در مقیاس بزرگ، تقسیم‌بندی عناصر، تشخیص نقطه کلیدی و «شرح‌نگاری تصویر»^۵ به کار می‌رود. نسخه‌ی اولیه‌ی این مجموعه، در سال 2014 حاوی 164 هزار تصویر بود که در نسخه‌ی 2017 به 328 هزار تصویر افزایش یافت که 123 هزار تصویر این نسخه، خام و بدون هرگونه نشانه‌گذاری است. در این داده، به‌ازای هر تصویر یک شرح وجود دارد.

مجموعه‌ی داده‌ی دیگر، توسط یونگ^۶ و همکاران (2014) تهیه شده‌است. این مجموعه، حاوی بیش از 31 هزار تصویر از وبگاه عکاسی فلیکر^۷ با موضوع فعالیت‌های روزانه‌ی افراد و فعالیت‌های اجتماعی است. این مجموعه‌ی داده Flickr30k نامیده می‌شود. ویژگی این مجموعه‌ی داده این است که به‌ازای هر تصویر، پنج شرح متفاوت نگارش شده‌است. در فرایند برچسب‌گذاری این مجموعه‌ی داده، از روش جمع‌سپاری استفاده شده‌است که نتیجه‌ی آن 158,915 توصیف برای تصاویر این مجموعه است.

توصیف تصویر دو مجموعه‌ی داده «عناصر مشترک در بافت» و فلیکر، توسط پژوهشگرانی، مانند نوید کنعانی و سجاد ایوبی، به فارسی ترجمه شده‌است و یک مجموعه‌ی داده‌ی ترکیبی از نظر حجم تصویر همراه با یک جمله‌ی شرح فارسی به‌دست آمده‌است.^۸ در این مجموعه‌ی داده، به‌ازای هر تصویر، یک شرح موجود است. در پژوهش دیگر، آرمان ملک‌زاده یک مجموعه‌ی داده‌ای با حجم بیش از 39 هزار تصویر را از خبرگزاری تسنیم گردآوری کرده‌است که حاوی خبر، تصویر مربوط به خبر و برچسب مقوله‌ای آن است.^۹

از میان مجموعه‌های داده‌ی موجود معرفی‌شده، مجموعه‌ی داده‌ی Flickr30k برای این پژوهش مناسب‌تر است؛ چراکه به‌ازای هر تصویر، پنج شرح متفاوت موجود است و این شیوه‌ی شرح تصویر، مفهوم «توجه» را در بر می‌گیرد؛ به این صورت که در هر شرح، فقط نشانه‌های مشخصی از نظام را شامل می‌شود و نشانه‌های مختلف با «توجه»های متفاوت در نظام تصویر، در نظام زبان جاری شده‌است. در شکل 3، نمونه‌ای از شیوه‌ی برچسب‌گذاری در مجموعه‌ی داده‌ی Flickr30k نمایش داده شده‌است.

محتوایی است که در شکل 2 قابل مشاهده است. شکل‌گیری این فضای انتزاعی، زیاد پیچیده نیست؛ چراکه براساس «توجه» به قسمتی از نظام تصویر در نظام زبان انتخاب می‌شود و این فضای انتزاعی براساس قالب‌های معنایی فریم‌نت که پیشتر مبتنی بر نمونه‌های زبانی به‌عنوان طرح‌واره‌های انتزاعی در ذهن گویش‌ور تعریف شده‌است شکل داده می‌شود. انتخاب این قالب‌های معنایی در این فضا، به نشانه‌های نظام تصویر و زبان وابسته است. هدف دیگر انجام این پژوهش، دسته‌بندی تصاویر براساس دانش انتزاعی حاصل از قالب‌های معنایی است. مراحل اجرای پژوهش در شکل 4 نمایش داده شده‌است.



شکل 4. مراحل روش‌شناسی پژوهش

همان‌طور که در شکل 4 مشخص است، مراحل اجرای پژوهش را به پنج مرحله تقسیم کرده‌ایم:

1. تهیه‌ی پیکره‌ی تصویری دارای شرح زبانی.
2. یافتن فعل اصلی در جمله‌ی شرح که کانون «توجه» در تصویر را شکل داده‌است و منجر به خلق این شرح شده‌است.
3. استخراج اطلاعات مربوط به نام قالب معنایی فعل از فریم‌نت^۱.
4. دسته‌بندی تصاویر براساس قالب معنایی مشترک افعال.
5. تبدیل دسته‌بندی به پایگاه دانش با استفاده از روابط معنایی میان تصاویر.

برای انجام این پژوهش داده‌محور، به تصویر دارای شرح نیاز است. چندین مجموعه‌ی داده با این ویژگی برای زبان انگلیسی تهیه شده‌است که به معرفی آنها می‌پردازیم. یک مجموعه‌ی داده نسبتاً پرجمعی که توسط شرکت مایکروسافت^۲ (لین و همکاران، 2014) تهیه شده‌است «عناصر مشترک در بافت»^۳ نام دارد.^۴ این مجموعه‌ی داده برای شناسایی عناصر

5. image captioning

6. Young, P.

7. <https://www.flickr.com/>

8. <https://www.kaggle.com/datasets/navidkanaani/coco-flickr-farsi>

9. <https://www.kaggle.com/datasets/malekzadeharman/persian-image-captioning-dataset>

1. <https://framenet.icsi.berkeley.edu/luIndex>

2. Microsoft

3. Common Objects in Context

4. <https://cocodataset.org/>

تحلیل داده

پژوهش حاضر باید در دو مرحله انجام پذیرد. در مرحله اول باید در نظام تصویر و زبان، مفهوم انتزاعی تصویر در چارچوب معناشناسی قالبی براساس نشانه‌های محتوایی زبانی از شرح تصاویر استخراج گردد؛ سپس براساس این عناصر معنایی به دسته‌بندی تصاویر اقدام گردد تا نقش‌های معنایی که می‌تواند در فضای انتزاعی میان نظام زبان و تصویر مطرح گردد مشخص شود. انجام این فعالیت‌ها با نظارت انسانی و به‌صورت دستی انجام می‌پذیرد که به بررسی نمونه‌هایی از داده‌های تحلیل‌شده می‌پردازیم.

در فرایند تحلیل داده‌ها چند قاعده مد نظر قرار گرفته‌است تا یکدستی در داده اعمال شود: (۱) بعضی از شرح‌های تصاویر مجموعه داده Flickr30k به‌صورت عبارت است. در این موارد، فعل «be» به‌عنوان فعل اصلی لحاظ می‌گردد. (۲) گاهی یک جمله به‌واسطه واو عطف بیش از یک فعل اصلی دارد. در این موارد، اولویت فعل اصلی براساس ترتیب حضور در جمله انتخاب می‌شود. چنانچه شواهد زبانی بر روی یک فعل از نظر بسامدی بیشتر باشد، این اولویت به فعل بعدی داده می‌شود.

دسته‌بندی تصاویر براساس قالب معنایی

در این بخش به بررسی نمونه‌هایی از داده‌های بررسی شده می‌پردازیم. شکل‌های ۵ تا ۱۰ نمونه‌های تصاویری است که به‌همراه شرح‌های نوشته‌شده مربوط به این تصاویر از مجموعه داده Flickr30k استخراج شده‌است و سپس قالب معنایی فعل جمله براساس «توجه» به قسمتی از تصویر فریم‌نت استخراج شده و به شرح هر تصویر تخصیص یافته‌است.







1	A man wearing a yellow safety vest helps prepare an airplane .
2	A man by a private jet is standing in the shade of the jet .
3	An aircraft stands on cement by workers in yellow vests .
4	An airplane is parked in front of a few vehicles .
5	A man is hanging out of the window of an airplane .


شکل ۵. نمونه‌ای از مجموعه داده Flickr30k همراه با پنج شرح مربوط به تصویر


«توجه» در جمله اول بر مرد جلیقه‌پوشی است که مشغول آماده‌سازی هواپیما برای پرواز است. وضعیت ایستادن این مرد در جمله دوم مورد «توجه» است. «توجه» در جمله سوم، به موقعیت هواپیما میان کارکنانی که روی زمین سیمانی ایستاده‌اند مربوط می‌شود. در جمله چهارم «توجه» بیانگر موقعیت هواپیما در مقایسه با سایر ماشین‌ها است؛ و «توجه» در جمله پنجم بر کمک‌خلبان متمرکز است. همان‌گونه که ملاحظه می‌گردد، «توجه» در هر یک از شرح‌های ارائه‌شده مربوط به تصویر، متفاوت از یکدیگر است و این تفاوت با به‌کارگیری نشانه‌های زبانی در نظام زبان منعکس شده‌است.


دسته اول

	1	Some women are standing in front of a bus with buildings behind it .	Being_located	1-5
	2	Several women stand on a city street with tall buildings .	Being_located	
	3	A group of women are standing in front of a bus .	Being_located	
	4	Several women wait outside in a city .	Waiting	
	5	Women are standing outside	Being_located	

	1	Someone in a blue shirt and hat is standing on stair and leaning against a window .	Being_located	2-5
	2	A man in a blue shirt is standing on a ladder cleaning a window .	Being_located	
	3	A man on a ladder cleans the window of a tall building .	Emptying	
	4	man in blue shirt and jeans on ladder cleaning windows	Emptying	
	5	a man on a ladder cleans a window	Emptying	
	1	Two men , one in a gray shirt , one in a black shirt , standing near a stove .	Being_located	3-5
	2	Two guy cooking and joking around with the camera .	Cooking_creation	
	3	Two men in a kitchen cooking food on a stove .	Cooking_creation	
	4	Two men are at the stove preparing food .	Cooking_creation	
	5	Two men are cooking a meal .	Cooking_creation	
	1	A young girl wearing a multicolored dress holds an orange ball in her right hand while walking through bright green grass behind a house.	Manipulation	4-5
	2	A little girl wearing a dress walks across a green lawn on a summer day with a home and trees in the background .	Self_motion	
	3	A child in a dress runs through the yard with a miniature orange soccer ball in her hand .	Self_motion	
	4	A little girl stands in a backyard , holding a toy and looking into the camera .	Being_located	
	5	This little girl is enjoying the spring weather after the long winter break .	Emotions_of_mental_acti vity	
	1	A group of adults , inside a home , sitting on chairs arranged in a circle , playing a type of musical instruments .	Cause_to_make_noise	5-5
	2	Five musicians , a man and four women , practicing sheet music (using flutes) in a living room .	Practice	
	3	People gathered in a circle , some holding musical instruments .	Come_together	
	4	People gathered in a room to talk about their favorite tunes .	Come_together	
	5	Five people are sitting in a circle with instruments .	Being_located	

	1	Two young guys with shaggy hair look at their hands while hanging out in the yard .	Perception_active	6-5
	2	Two young , White males are outside near many bushes.	Performers_and_roles	
	3	Two men in green shirts are standing in a yard .	Being_located	
	4	A man in a blue shirt standing in a garden .	Being_located	
	5	Two friends enjoy time spent together .	Emotions_of_mental_activity	

	1	A man sits in a chair while holding a large stuffed animal of a lion .	Being_located	7-5
	2	A man is sitting on a chair holding a large stuffed animal .	Being_located	
	3	A man completes the finishing touches on a stuffed lion.	Activity_finish	
	4	A man holds a large stuffed lion toy .	Manipulation	
	5	A man is smiling at a stuffed lion	Making_faces	





	1	An asian man wearing a black suit stands near a dark-haired woman and a brown-haired woman .	Being_located	8-5
	2	Three people are standing outside near large pipes and a metal railing .	Being_located	
	3	A young woman walks past two young people dressed in hip black outfits .	Self_motion	
	4	A woman with a large purse is walking by a gate .	Self_motion	
	5	Several people standing outside a building.	Being_located	

شکل ۵. نمونه‌هایی از تصویر و قالب معنایی مشترک «Being_located»

رفته‌است، در این قالب معنایی قرار می‌گیرد. بنابراین این ویژگی معنایی به‌عنوان دانش انتزاعی، براساس «توجه» به قسمتی از تصاویر در موقعیتی اشاره دارد که با نشان‌های زبانی «حضورداشتن»، «نشستن» یا «ایستادن» به‌عنوان نقطه اشتراک میان این ۸ تصویر شده‌است. با استفاده از این ویژگی مشترک می‌توان به دسته‌بندی تصاویری پرداخت که ممکن است در نگاه اول اشتراکی میان آن‌ها دیده نشود.

با بررسی شرح‌های ارائه‌شده در شکل ۵ برای هر تصویر و همچنین قالب‌های معنایی تخصیص‌یافته به هر شرح براساس فعل اصلی جمله می‌توان دریافت که قالب معنایی «Being_located» (قرارداشتن) در میان این مجموعه از تصاویر، مشترک است که برجسب‌های مد نظر برجسته شده‌است. با بررسی فعل‌های جملات شرح، مشخص شد که افعالی مانند «be»، «sit» و «stand» ازجمله افعالی است که باتوجه به بافت زبانی که در شرح‌های این تصاویر به‌کار

دسته دوم



	1	Three young men and a young woman wearing sneakers are leaping in midair at the top of a flight of concrete stairs .	Self_motion	1-6
	2	Four casually dressed guys jumping down a stairway outdoors with a stone wall behind them .	Change_position_on_a_scale	
	3	Four guys three wearing hats one not are jumping at the top of a staircase .	Change_position_on_a_scale	
	4	Four men with excited faces are jumping down from the top of stairs .	Change_position_on_a_scale	
	5	Four people are jumping from the top of a flight of stairs .	Change_position_on_a_scale	
	1	Two men in Germany jumping over a rail at the same time without shirts .	Change_position_on_a_scale	2-6
	2	Two youths are jumping over a roadside railing , at night .	Change_position_on_a_scale	
	3	Boys dancing on poles in the middle of the night .	Self_motion	
	4	Two men with no shirts jumping over a rail .	Change_position_on_a_scale	
	5	two guys jumping over a gate together	Change_position_on_a_scale	
	1	Five ballet dancers caught mid jump in a dancing studio with sunlight coming through a window .	Come_down_with	3-6
	2	Ballet dancers in a studio practice jumping with wonderful form .	Practice	
	3	Five girls are leaping simultaneously in a dance practice room .	Self_motion	
	4	Five girls dancing and bending feet in ballet class .	Self_motion	
	5	A ballet class of five girls jumping in sequence .	Change_position_on_a_scale	
	1	A child in a pink dress is climbing up a set of stairs in an entry way .	Change_position_on_a_scale	4-65
	2	A little girl in a pink dress going into a wooden cabin .	Motion	
	3	A little girl climbing the stairs to her playhouse .	Change_position_on_a_scale	
	4	A little girl climbing into a wooden playhouse	Change_position_on_a_scale	
	5	A girl going into a wooden building .	Motion	


شکل ۶. نمونه‌هایی از تصویر و قالب معنایی مشترک «Change_position_on_a_scale»

سبب دسته‌بندی این مجموعه از تصاویر در یک دسته شده‌است که نمونه‌های مرتبط با این قالب برجسته شده‌است. مفهوم «تغییر موقعیت» ویژگی معنایی مشترکی است که در نشانه‌های زبانی مانند «پريدن» و «بالارفتن» متبلور بوده، کانون توجه در این تصاویر است.

با بررسی شرح‌های ارائه‌شده در شکل ۶ و قالب‌های معنایی تخصیص‌یافته به هر شرح براساس فعل اصلی جمله این نتیجه به‌دست آمد که قالب معنایی «Change_position_on_a_scale» (تغییر موقعیت در مقیاس) در افعالی مانند «jump» و «climb» وجود دارد که

دسته سوم



	1	A young woman with dark hair and wearing glasses is putting white powder on a cake using a sifter .	Placing	1-7
	2	A lady in a black top with glasses is sprinkling powdered sugar on a bundt cake .	Filling	
	3	A woman wearing glasses sprinkles powdered sugar onto her bundt cake .	Filling	
	4	Girl in black jacket sifting powdered sugar over a chocolate cake .	Scouring	
	5	A standing woman holds a pan over a cake .	Manipulation	
	1	Two people in the photo are playing the guitar and the other is poking at him .	Make_noise	2-7
	2	A man in green holds a guitar while the other man observes his shirt .	Manipulation	
	3	A man is fixing the guitar players costume .	Activity_prepare	
	4	a guy stitching up another man 's coat .	Attaching	
	5	the two boys playing guitar	Make_noise	
	1	a man sleeping on a bench outside with a white and black dog sitting next to him .	Sleep	3-7
	2	A man lays on the bench to which a white dog is also tied .	Placing	
	3	man laying on bench holding leash of dog sitting on ground	Manipulation	
	4	A shirtless man lies on a park bench with his dog .	Placing	
	5	A man lays on a bench while his dog sits by him .	Placing	
	1	A young girl wearing a multicolored dress holds an orange ball in her right hand while walking through bright green grass behind a house.	Manipulation	4-7
	2	A little girl wearing a dress walks across a green lawn on a summer day with a home and trees in the background .	Self_motion	
	3	A child in a dress runs through the yard with a miniature orange soccer	Self_motion	




		ball in her hand .		
		A little girl stands in a backyard , holding a toy and looking into the camera .	Being_located	
		This little girl is enjoying the spring weather after the long winter break .	Emotions_of_mental_activity	
	1	A man sits in a chair while holding a large stuffed animal of a lion .	Being_located	5-7
	2	A man is sitting on a chair holding a large stuffed animal .	Being_located	
	3	A man completes the finishing touches on a stuffed lion .	Activity_finish	
	4	A man holds a large stuffed lion toy .	Manipulation	
	5	A man is smiling at a stuffed lion	Making_faces	

شکل ۷. نمونه‌هایی از تصویر و قالب معنایی مشترک «Manipulation»

شده‌است که نمونه‌های مرتبط با این قالب معنایی برجسته شده‌است. وجه مشترک این تصاویر «نگهداری چیزی با دست» است که با نشانه‌ی زبانی «نگه‌داشتن» تجلی یافته و کانون «توجه» در شرح تصاویر بوده‌است.

از بررسی شرح‌های ارائه‌شده در شکل ۷ و قالب‌های معنایی تخصیص‌یافته به هر شرح براساس فعل اصلی جمله این نتیجه به دست آمد که قالب معنایی «Manipulation» (انجام دادن با دست) در فعلی مانند «hold» وجود دارد که سبب دسته‌بندی این مجموعه از تصاویر در یک دسته‌ی چهارم

	1	Three young men and a young woman wearing sneakers are leaping in midair at the top of a flight of concrete stairs .	Self_motion	1-8
	2	Four casually dressed guys jumping down a stairway outdoors with a stone wall behind them .	Change_position_on_a_scale	
	3	Four guys three wearing hats one not are jumping at the top of a staircase .	Change_position_on_a_scale	
	4	Four men with excited faces are jumping down from the top of stairs .	Change_position_on_a_scale	
	5	Four people are jumping from the top of a flight of stairs .	Change_position_on_a_scale	
	1	Two men in Germany jumping over a rail at the same time without shirts .	Change_position_on_a_scale	2-8
	2	Two youths are jumping over a roadside railing , at night .	Change_position_on_a_scale	
	3	Boys dancing on poles in the middle of the night .	Self_motion	
	4	Two men with no shirts jumping over a rail .	Change_position_on_a_scale	
	5	two guys jumping over a gate together	Change_position_on_a_scale	




	1	Five ballet dancers caught mid jump in a dancing studio with sunlight coming through a window .	Come_down_with	3-8
	2	Ballet dancers in a studio practice jumping with wonderful form .	Practice	
	3	Five girls are leaping simultaneously in a dance practice room .	Self_motion	
	4	Five girls dancing and bending feet in ballet class .	Self_motion	
	5	A ballet class of five girls jumping in sequence .	Change_position_on_a_scale	
	1	A young girl wearing a multicolored dress holds an orange ball in her right hand while walking through bright green grass behind a house.	Manipulation	4-8
	2	A little girl wearing a dress walks across a green lawn on a summer day with a home and trees in the background .	Self_motion	
	3	A child in a dress runs through the yard with a miniature orange soccer ball in her hand .	Self_motion	
	4	A little girl stands in a backyard , holding a toy and looking into the camera .	Being_located	
	5	This little girl is enjoying the spring weather after the long winter break .	Emotions_of_mental_activity	
	1	An asian man wearing a black suit stands near a dark-haired woman and a brown-haired woman .	Being_located	5-8
	2	Three people are standing outside near large pipes and a metal railing .	Being_located	
	3	A young woman walks past two young people dressed in hip black outfits .	Self_motion	
	4	A woman with a large purse is walking by a gate .	Self_motion	
	5	Several people standing outside a building .	Being_located	

شکل ۸. نمونه‌هایی از تصویر و قالب معنایی مشترک «Self_motion»

تصاویر در یک دسته شده‌است که نمونه‌های با این قالب معنایی برجسته شده‌است. مفهوم «خود حرکتی» ویژگی معنایی مشترکی است که در نشانه‌های زبانی مانند «جستن»، «رقصیدن» و «پیاده رفتن» متبلور بوده، کانون «توجه» در این تصاویر است.

بررسی شرح‌های ارائه‌شده در شکل ۸ و قالب‌های معنایی تخصیص‌یافته به هر شرح براساس فعل اصلی جمله منجر به این نتیجه شد که قالب معنایی «Self_motion» (خود حرکتی) در افعالی مانند «leap»، «dance» و «walk» وجود دارد که سبب دسته‌بندی این مجموعه از

دسته پنجم


	1	Several men in hard hats are operating a giant pulley system .	Being_in_operation	1-9
	2	Workers look down from up above on a piece of equipment .	Perception_active	
	3	Two men working on a machine wearing hard hats .	Work	
	4	Four men on top of a tall structure .	Performers_and_roles	
	5	Three men on a large rig .	Performers_and_roles	
	1	A black dog and a white dog with brown spots are staring at each other in the street .	Perception_active	2-9
	2	A black dog and a tri-colored dog playing with each other on the road .	Competition	
	3	Two dogs of different breeds looking at each other on the road .	Perception_active	
	4	Two dogs on pavement moving toward each other .	Motion	
	5	A black dog and a spotted dog are fighting	Hostile_encounter	
	1	Two young guys with shaggy hair look at their hands while hanging out in the yard .	Perception_active	3-9
	2	Two young , White males are outside near many bushes .	Performers_and_roles	
	3	Two men in green shirts are standing in a yard .	Being_located	
	4	A man in a blue shirt standing in a garden .	Being_located	
	5	Two friends enjoy time spent together .	Emotions_of_mental_activity	


شکل ۸. نمونه‌هایی از تصویر و قالب معنایی مشترک «Perception_active»

یک دسته شده‌است که نمونه‌های ارائه‌شده مرتبط با این قالب معنایی در هر شرح برجسته شده‌است. وجه مشترک این تصاویر «نگاه» است که با نشانه‌ی زبانی «دیدن» و «خیره شدن» تجلی یافته، کانون «توجه» در شرح تصاویر بوده‌است.

قالب معنایی «Perception_active» (ادراک فعال) از بررسی شرح‌های ارائه‌شده در شکل 9 و قالب‌های معنایی مشترک تخصیص‌یافته به هر شرح براساس فعل اصلی جمله تعیین شده‌است که در افعالی مانند «look down»، «stare»، «look at» و «look» وجود دارد. این مفهوم انتزاعی مشترک سبب دسته‌بندی این مجموعه از تصاویر در

دسته ششم

	1	A young woman with dark hair and wearing glasses is putting white powder on a cake using a sifter .	Placing	1-10
	2	A lady in a black top with glasses is sprinkling powdered sugar on a bundt cake .	Filling	
	3	A woman wearing glasses sprinkles powdered sugar onto her bundt cake .	Filling	
	4	Girl in black jacket sifting powdered sugar over a chocolate cake .	Scouring	
	5	A standing woman holds a pan over a cake .	Manipulation	

	1	a man sleeping on a bench outside with a white and black dog sitting next to him .	Sleep	2-10
	2	A man lays on the bench to which a white dog is also tied .	Placing	
	3	man laying on bench holding leash of dog sitting on ground	Manipulation	
	4	A shirtless man lies on a park bench with his dog .	Placing	
	5	A man lays on a bench while his dog sits by him .	Placing	

شکل ۱۰. نمونه‌هایی از تصویر و قالب معنایی مشترک «Placing»

با نگاهی به شش دسته داده بررسی شده مشخص شد که قالب معنایی مربوط به فعل اصلی جمله که براساس روش «توجه» با نشانه‌های زبانی به قسمتی از تصویر اشاره می‌کند به چه صورت می‌تواند در دسته‌بندی نشانه‌های تصویری نقش داشته باشد و مورد استفاده قرار گیرد. با نگاهی به دسته‌بندی‌ها انجام شده این نکته جلب توجه نمود که بعضی از تصاویر در بیش از یک دسته قرار گرفته‌است که به نمونه‌هایی از این موارد که در جدول ۲ گردآوری شده‌است، اشاره می‌شود.

با بررسی شرح‌های ارائه شده در شکل ۷ و قالب‌های معنایی تخصیص یافته به هر شرح براساس فعل اصلی جمله این نتیجه به دست آمد که قالب معنایی «Placing» (قرار گرفتن) در افعالی مانند «put» و «lay» وجود دارد که سبب دسته‌بندی این مجموعه از تصاویر در یک دسته شده‌است که نمونه‌های ارائه شده مرتبط با این قالب برجسته شده‌است. وجه مشترک این تصاویر «وضعیتی در قرار گرفتن» است که با نشانه زبانی «قرار دادن» یا «دراز کشیدن» تجلی یافته و کانون «توجه» در شرح تصاویر بوده‌است.

جدول ۲. نمونه‌های تصاویر دسته‌بندی شده در بیش از یک دسته براساس قالب معنایی شرح تصویر

ارجاع دسته نمونه	نام قالب معنایی	تصویر	ردیف
6-5	Being_located (قرارداشتن)		۱
3-9	Perception_active (ادراک فعال)		
7-5	Being_located (قرارداشتن)		۲
5-7	Manipulation (انجام دادن با دست)		
8-5	Being_located (قرارداشتن)		۳

ارجاع دسته نمونه	نام قالب معنایی	تصویر	ردیف
5-8	Self_motion (خود حرکتی)		
1-6	Change_position_on_a_scale (تغییر موقعیت در مقیاس)		۴
1-8	Self_motion (خود حرکتی)		
2-6	Change_position_on_a_scale (تغییر موقعیت در مقیاس)		۵
2-8	Self_motion (خود حرکتی)		
3-6	Change_position_on_a_scale (تغییر موقعیت در مقیاس)		۶
3-8	Self_motion (خود حرکتی)		
1-7	Manipulation (انجام دادن با دست)		۷
1-10	Placing (قرار گرفتن)		
3-7	Manipulation (انجام دادن با دست)		۸
2-10	Placing (قرار گرفتن)		
4-5	Being_located (قرار داشتن)		9
4-7	Manipulation (انجام دادن با دست)		
4-8	Self_motion (خود حرکتی)		

چند تصویر، مانند ردیف‌های ۴ تا ۶ جدول ۲ که در آنها تحرک و تغییر موقعیت مفهوم انتزاعی مشترک در نشانه‌های تصویری

از تصاویر گردآوری شده در جدول ۲ می‌توان به دو جمع‌بندی رسید: الف) وجود یک مفهوم انتزاعی میان یک یا

است، با افعال مختلفی به‌عنوان نشانه‌های زبانی بیان شده‌است؛ ب) وجود مفاهیم انتزاعی مختلف در یک تصویر، مانند موارد گزارش شده در جدول ۲، براساس روش «توجه» در نشانه‌های تصویری در قالب جملات متنوع در نشانه‌های زبانی بیان شده‌است. ردیف ۹ این جدول یکی از مواردی است که مفاهیم انتزاعی مختلفی را در خود جای داده‌است.

در بخش ۲-۴ توضیحات کلی درمورد فریم‌نت ارائه شد. در آن بخش به این نکته اشاره شد که هر قالب معنایی، تعدادی عناصر قالبی دارد که همچون طرح‌واره عمل می‌کند. این عناصر ممکن است در جمله موجود باشد یا نباشد. در جدول ۳ عناصر هر قالب که از فریم‌نت انگیزی استخراج شده، فهرست شده‌است.

عناصر معنایی مربوط به قالب معنایی

جدول ۳. فهرست قالب‌های مطرح‌شده در تصاویر و عناصر مربوط به قالب‌ها

عناصر قالب	نام قالب
Location, Place, Theme	Being_located (قرارداشتن)
Attribute, Circumstances, Correlated_variable, Difference, Final_correlate, Final_value, Initial_value, Item, Manner, Result, Time, Value_range	Change_position_on_a_scale (تغییر موقعیت در مقیاس)
Agent, Entity, Locus	Manipulation (انجام دادن با دست)
Area, Concessive, Cotheme, Depictive, Direction, Distance, Duration, Explanation, Goal, Internal_cause, Manner, Means, Path, Place, Purpose, Result, Self_mover, Source, Speed, Time	Self_motion (خود حرکتی)
Body_part, Depictive, Direction, Location_of_protagonist, Manner, Perceiver_agentive, Phenomenon, State	Perception_active (ادراک فعال)
Agent, Cause, Duration, Explanation, Goal, Manner, Theme, Time	Placing (قرار گرفتن)

است که این عناصر معنایی برای تمامی قالب‌های معنایی بدون هرگونه کم‌وکاستی قابلیت استفاده دارد و ثابت است؛ درحالی‌که ممکن است عناصر قالب با توجه به نوع قالب و معنای فعل مورد نظر، متفاوت باشد. عناصر قالب معرفی شده سبب ایجاد نوعی یکدستی در ساختار تهیه داده می‌شود.

الف: **کنش**: منظور از کنش، اتفاق یا رویدادی است که در تصویر افتاده‌است. این نقش معنایی معمولاً به‌صورت ضمنی از ارتباط میان نشانه‌ها در نظام تصویر یا شرح تصویر در نظام زبان به‌دست می‌آید. معمولاً یک کنش توسط فرد یا افرادی بر روی یک پدیده‌ای تحقق می‌پذیرد، همانند جمله معلوم در نظام زبان. این امکان وجود دارد که کنش‌گر در تصویر حضور نداشته باشد ولی اثر کنش‌گر در نظام تصویر قابل مشاهده باشد، همانند جمله مجهول در نظام زبان. برای

از آنجاکه در این پژوهش با تصویر و متن مرتبط با تصویر سروکار داریم، می‌توانیم علاوه بر اطلاعات مربوط به عناصر قالب از متن، عناصر معنایی که طبق نظر شعیری (۱۳۹۷) در چارچوب نشانه-معناشناسی دیداری امکان استخراج از تصویر دارد را به مجموعه اطلاعات گردآوری شده اضافه کنیم و از این اطلاعات در قالب فراداده برای ساخت یک پایگاه دانش استفاده نماییم. ولی در این پژوهش، خود را به اطلاعات معنایی انتزاعی حاصل از قالب‌های معنایی استخراج‌شده از متن‌های مرتبط با تصاویر محدود کرده‌ایم. ممکن است عناصر معنایی حاصل از تصویر که توسط شعیری (۱۳۹۷) معرفی شده‌است با بعضی از عناصر قالب‌های معنایی مشترک باشد. تفاوت عناصر قالب معرفی شده در جدول ۳ و عناصر معنایی که در ادامه این بخش به معرفی آنها می‌پردازیم این

بخش‌های زمانی یک روز، مانند صبح، روز، شب، عصر، هوای تاریک، هوای گرگ‌ومیش و مانند آن، است به نشانه‌ی موجود در نظام تصویر مربوط است. این نقش معنایی ممکن است در جمله به صورت گروه حرف اضافه‌ای یا قید در نظام زبان مطرح گردد.

ج: ابزار کنش: منظور از ابزار کنش ابزاری است که ممکن است کنش‌گر برای کنش خود از آن استفاده کند. این نقش معنایی به صورت گروه حرف اضافه‌ای که هسته‌ی آن حرف اضافه‌ی «با» است در نظام زبان مطرح می‌گردد.

چ: وضعیت کنش‌گر: منظور از وضعیت کنش‌گر حالتی است که کنش‌گر در آن مکان کنش، به کنش‌گری می‌پردازد، برای مثال، سر میز غذا، روی پله، روی نردبان و مانند آن، که به توصیف وضعیت کنش‌گر در نظام زبان می‌پردازد. این نقش معنایی نباید با مکان یا صحنه‌ی کنش اشتباه گردد. این نقش معنایی ممکن است ضمنی بوده و براساس عناصر موجود در تصویر قابل گمانه‌زنی باشد.

ح: احساس کنش‌گر: منظور از احساس کنش‌گر هیچانی است که شرکت‌کننده، در کنش تصویر از خود به نمایش گذاشته و به بیننده منتقل می‌کند. چنانچه کنش‌گر موجود زنده انسانی باشد، هیجانات احساسی وی از روی صورت یا حالت بدن قابل ادراک است.

خ: هدف کنش‌گر: منظور از هدف کنش‌گر هدفی است که کنش‌گر به انجام آن کنش می‌پردازد. این نقش معنایی معمولاً ضمنی بوده و براساس سایر عناصر موجود در تصویر قابل گمانه‌زنی است، برای مثال، هدف از استفاده از قایق، مسافرت یا تفریح است.

در ادامه، برای نمونه، روش تحلیل داده‌ها در تعیین و دسته‌بندی عناصر معنایی معرفی شده در نظام تصویر و نظام زبان را بر شکل‌های ۳ و ۴-۴ اعمال می‌کنیم؛ چراکه در این شکل‌ها براساس «توجه»، برای یک نظام تصویر پنج شرح متفاوت در نظام زبان نگارش شده است و عناصر معنایی هر یک از نشانه‌ها در نظام زبان با نشانه‌ها در نظام تصویر متفاوت است. در جدول ۴، نقش‌های معنایی عناصر موجود در نظام زبان و تصویر شکل ۳ گزارش شده است.

مثال در جمله «شیشه شکست»، کنش‌گری که سبب شکستن شیشه شده است نامشخص است که می‌تواند فردی باشد که با وسیله‌ای اقدام به شکستن شیشه کرده باشد. در مثالی دیگر، «توپ شیشه را شکست»، اگرچه در نظام زبان، کنش «شکستن» توسط «توپ» بر روی «شیشه» اعمال شده است، این «توپ» توسط کنش‌گری که مستتر است به حرکت در آمده است. بنابراین، نوع کنش در نظام تصویر همچون نظام زبان باید مشخص شود که کنش توسط کنش‌گر انجام شده است یا فقط اثر کنش در تصویر قابل مشاهده است. از آنجاکه ماهیت تصویر به تحقق یک رویداد یا کنش در زمان گذشته اشاره دارد، گاهی وضعیت کنش به صورت استمراری در نظام تصویر قابل مشاهده است. بنابراین وضعیت کنش «انجام‌یافته» یا «در حال انجام» باید مشخص گردد.

ب: کنش‌گر: منظور از کنش‌گر شرکت‌کننده‌ای است که به کنش در تصویر می‌پردازد. کنش‌گر می‌تواند جاندار یا شیء باشد. یکی از ویژگی‌های مربوط به کنش‌گر، تعداد کنش‌گر است که می‌تواند مفرد یا جمع باشد. ویژگی دیگر مربوط به کنش‌گر، جنسیت ژنتیکی است که می‌تواند مذکر، مؤنث یا خنثی باشد. چنانچه شرکت‌کننده‌ی کنش یک نفر باشد، جنسیت یکی از سه حالت ذکر شده است؛ ولی اگر چند نفر باشد، می‌تواند یکی از سه جنسیت ذکر شده یا به صورت گروهی و ترکیبی از همه جنسیت‌ها باشد.

پ: کنش‌پذیر: منظور از کنش‌پذیر جاندار یا شیئی است که به کنش مرتبط است و کنش بر آن واقع می‌شود.

ت: مکان یا صحنه‌ی کنش: منظور از مکان یا صحنه‌ی کنش، موقعیت مکانی است که کنش در آن اتفاق می‌افتد که در نظام تصویر قابل مشاهده است. این مکان به یک فضا، مانند اتاق، خیابان، حیاط، مزرعه و غیره، مربوط است و نه روی میز، روی نردبان و مانند آن. چراکه میز و نردبان خودشان از جمله عناصری است که باید در یک فضا، مانند اتاق، قرار بگیرد. این نقش معنایی در نظام زبان به صورت گروه حرف اضافه‌ای که هسته‌ی آن حرف اضافه‌ی «در» است مطرح می‌گردد.

ث: زمان کنش: منظور از زمان کنش موقعیت زمانی است که کنش در آن اتفاق می‌افتد. این زمان که مربوط به

جدول ۴. نقش‌های معنایی عناصر موجود در نظام زبان و تصویر شکل ۳

نقش معنایی	جمله ۱	جمله ۲	جمله ۳	جمله ۴	جمله ۵
کنش	help, progressive, active	stand, progressive, active	stand, progressive, active	park, present, passive	hang out, progressive, active
کنش‌گر	man, single, male	man, single, male	aircraft, single, neutral	-	man, single, male
کنش‌پذیر	-	-	-	airplane	-
مکان کنش	<u>Airport</u>	<u>airport</u>	<u>airport</u>	<u>airport</u>	<u>airport</u>
زمان کنش	<u>Day</u>	<u>Day</u>	<u>day</u>	<u>day</u>	<u>day</u>
ابزار کنش	-	-	-	-	-
وضعیت کنش‌گر	<u>Standing</u>	standing	on cement	in front of vehicles	in the airplane
احساس کنش‌گر	-	-	-	-	-
هدف کنش‌گر	<u>responsibility</u>	<u>working</u>	<u>park</u>	park	<u>communication</u>

استفاده شده‌است و در دو جمله از فعل متعدی. کنش‌گر تمام جملات یک دختر بچه است. کنش‌پذیر در جملات اول و پنجم، نشانه‌های «توپ» و «هوا» است که با کنش‌های «نگه‌داشتن» و «لذت بردن» مرتبط است. مکان کنش در جملات اول تا چهارم مربوط به «پشت خانه»، «چمنزار»، «حیاط» و «حیاط پشتی» به صورت صریح مشخص است؛ در حالی که مکان کنش در جمله پنجم «فضای باز» است که در نظام زبان مطرح نشده و ضمنی است. زمان کنش جزء اطلاعات نامشخص است. ابزار مربوط به کنش وجود ندارد. وضعیت کنش‌گر در جملات اول و پنجم ضمنی است که در نظام زبان نامشخص است ولی در نظام تصویر مشخص است؛ در حالی که، در سایر جملات وضعیت کنش‌گر در نظام زبان به صورت صریح مشخص شده‌است. هدف کنش‌گر از انجام کنش معمولاً به صورت ضمنی قابل درک است که با توجه به تصویر، اهدافی مانند «مسئولیت‌شناسی» در جمله اول، «کار کردن» در جمله دوم، «پارک کردن» در جمله سوم و «ارتباط» در جمله پنجم از تصاویر قابل درک است. هدف کنش‌گر در جمله چهارم «پارک کردن» است که به صورت صریح در جمله به آن اشاره شده‌است.

برای کنش چهار جمله از میان پنج شرح ارائه شده در شکل ۳، از فعل لازم استفاده شده‌است و در یک شرح از فعل متعدی استفاده شده‌است که به صورت جمله مجهول به کار رفته‌است. کنش‌گر در جملات معلوم و کنش‌پذیر در جمله مجهول مشخص است. مکان و زمان کنش جزء اطلاعات ضمنی است که در نظام تصویر مشخص و در نظام زبان نامشخص است. ابزار مربوط به کنش وجود ندارد. وضعیت کنش‌گر در جملات اول و پنجم ضمنی است که در نظام زبان نامشخص است ولی در نظام تصویر مشخص است؛ در حالی که، در سایر جملات وضعیت کنش‌گر در نظام زبان به صورت صریح مشخص شده‌است. هدف کنش‌گر از انجام کنش معمولاً به صورت ضمنی قابل درک است که با توجه به تصویر، اهدافی مانند «مسئولیت‌شناسی» در جمله اول، «کار کردن» در جمله دوم، «پارک کردن» در جمله سوم و «ارتباط» در جمله پنجم از تصاویر قابل درک است. هدف کنش‌گر در جمله چهارم «پارک کردن» است که به صورت صریح در جمله به آن اشاره شده‌است.

در جدول ۵، نقش‌های معنایی عناصر موجود در نظام زبان و تصویر شکل ۴-۵ گزارش شده‌است. برای کنش سه جمله از میان پنج شرح ارائه شده در شکل ۴-۵، از فعل لازم

جدول ۵. نقش‌های معنایی عناصر موجود در نظام زبان و تصویر شکل ۴-۵

نقش معنایی	جمله ۱	جمله ۲	جمله ۳	جمله ۴	جمله ۵
کنش	hold, progressive, active	walk, progressive, active	run, progressive, active	stand, progressive, active	enjoy, progressive, active
کنش‌گر	girl, single, female	girl, single, female	child, single, female	girl, single, female	girl, single, female
کنش‌پذیر	Ball	-	-	-	weather
مکان کنش	behind the house	Lawn	yard	backyard	outside
زمان کنش	Day	Day	day	Day	day
ابزار کنش	Hand	Foot	foot	Foot	-
وضعیت کنش‌گر	Standing	standing	standing	Standing	standing
احساس کنش‌گر	delight	delight	delight	delight	delight
هدف کنش‌گر	amusement	amusement	amusement	amusement	amusement

مفاهیم براساس یک سه‌تایی^۱ بیان می‌شود به این مفهوم که میان دو عنصر یا موجودیت نوعی رابطه وجود دارد و این رابطه چیزی نیست جز مفهوم انتزاعی استخراج‌شده از شرح تصاویر. برای مثال در پایگاه دانش، اطلاعات به صورت سه‌تایی زیر بازنمایی می‌شود:

$X(Y,Z)$

در این مثال، Y و Z دو عنصر بوده و X بیانگر رابطه موجود میان دو عنصر Y و Z است. ویژگی این نوع سازماندهی اطلاعات این است که می‌توان با استفاده از زبان‌های جستجوی منطبق با منطق محمولات، به جستجوی تصاویری که به یک مفهوم انتزاعی مشخص اشاره دارد پرداخت. چنانچه شیوه سازمان‌دهی سه‌تایی اطلاعات را مد نظر دهیم و آن را با فعالیت مورد نظر خود منطبق کنیم، سه‌تایی که انتظار داریم از نشانه‌های تصویری و زبانی ساخته شود به صورت الگوی زیر خواهد بود:

(تصویر ۲، تصویر ۱) قالب معنایی

نکته قابل توجهی که می‌توان از مقایسه جداول نقش‌های معنایی ۴ و ۵ در مورد نقش‌های معنایی عناصر موجود در نظام زبان و تصویر به آن دست یافت این است بعضی از نقش‌های معنایی ضروری و بعضی دیگر فرعی است. نقش‌های معنایی ضروری می‌تواند موارد مانند کنش، کنش‌گر، مکان کنش و وضعیت کنش‌گر باشد و سایر نقش‌های معنایی به‌عنوان نقش معنایی فرعی تلقی شود. دلیل انتخاب این چهار نقش معنایی به‌عنوان نقش معنایی اصلی، وجود اطلاعات صریح در اکثر موارد و نه اطلاعات ضمنی در نشانه‌های زبانی و تصویری است.

تهیه پایگاه دانش از تصاویر و قالب‌های معنایی

در بخش ۵-۱ شیوه تحلیل داده‌ها در یافتن مفاهیم انتزاعی و دسته‌بندی تصاویر براساس مفهوم انتزاعی منعکس‌شده در قالب معنایی توضیح داده شد. در این بخش به چگونگی سازماندهی این اطلاعات در قالب تهیه پایگاه دانش می‌پردازیم. یکی از ویژگی‌های پایگاه دانش این است که

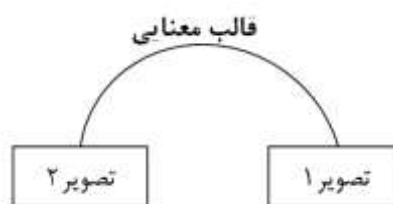
1. triple

Manipulation(5-4,7-1)
 Manipulation(5-4,7-2)
 Manipulation(5-4,7-3)
 Manipulation(5-4,7-4)

Self_motion(5-4,8-1)
 Self_motion(5-4,8-2)
 Self_motion(5-4,8-3)
 Self_motion(5-4,8-5)

روابط منطقی بالا به صورت مصور در شکل ۱۱ قابل مشاهده است. براساس روابط منطقی تعریف‌شده بالا، دو رابطه Being_located (قرارداشتن) و Manipulation (انجام دادن با دست) میان تصاویر ۴-۴ و ۷-۴ وجود دارد. همچنین دو رابطه Being_located (قرارداشتن) و Self_motion (خود حرکتی) میان تصاویر ۴-۴ و ۸-۴ وجود دارد. سازماندهی اطلاعات به این صورت می‌تواند بیانگر وجود نشانه‌های تصویری و زبانی مشترک میان دو تصویر و جمله باشد که گاهی ممکن است این رابطه در نظر اول وجود نداشته باشد. ولی با به‌کارگیری مفاهیم انتزاعی می‌توان به این کلیت دست یافت.

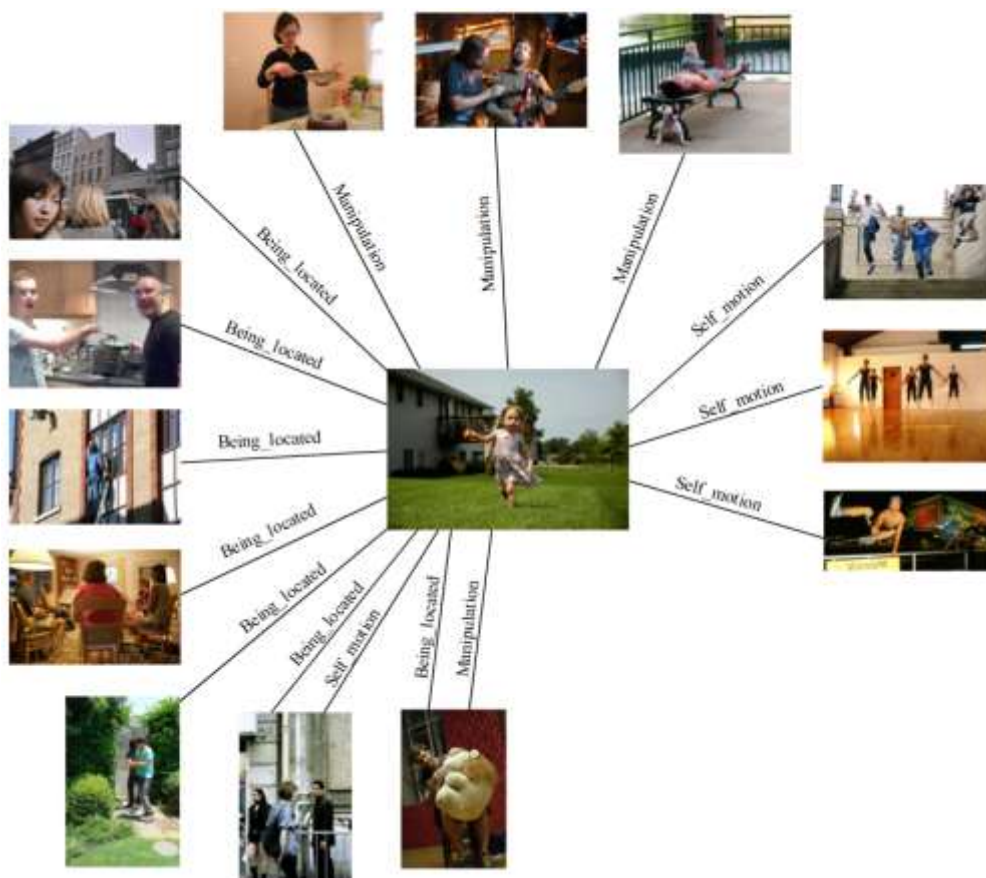
این سه‌تایی به صورت گرافیکی به صورت الگوی ترسیم‌شده در شکل ۱۰ بیان می‌گردد:



شکل ۱۰: الگوی سازماندهی تصاویر براساس قالب معنایی

براساس این الگو، سه‌تایی‌های زیر با محوریت تصویر ۵-۴ و شرح‌های نوشته‌شده مرتبط به این تصویر توسط گویشوران براساس روش «توجه» به دست خواهد آمد:

- +Being_located(5-4,5-1)
- Being_located(5-4,5-2)
- Being_located(5-4,5-3)
- Being_located(5-4,5-5)
- Being_located(5-4,5-6)
- Being_located(5-4,5-7)
- Being_located(5-4,5-8)



شکل ۱۱. پایگاه دانش حاصل از به‌کارگیری نام قالب‌ها به‌عنوان رابطه میان نشانه‌های تصویری

نتیجه‌گیری و جمع‌بندی

در پژوهش حاضر، در چارچوب نشانه‌شناسی رایانشی با رویکرد چندوجهی به مسئله‌ی صورت و معنا، به ایجاد ارتباط میان نشانه‌های زبانی و نشانه‌های تصویری پرداختیم. در راستای اهداف پژوهش، مفهوم انتزاعی موجود در متنی که براساس «توجه» به قسمتی از تصویر نگارش شده بود را از شبکه‌ی معنایی قالبی استخراج کردیم و با استفاده از آن به دسته‌بندی تصاویر و یافتن ارتباط معنایی میان تصاویر پرداختیم. نتایج حاصل از پژوهش از این قرار است:

الف) می‌توان با «توجه» به قسمت‌های مختلف تصویر، به توضیح و بیان عناصر و ویژگی‌های هر تصویر پرداخت. بنابراین، شرح هر تصویر الزاماً فقط یک جمله نیست.

ب) هر شرح تصویر، یک مفهوم انتزاعی و کلی در خود دارد که می‌تواند به‌عنوان معیار جهت یافتن ارتباط معنایی میان تصاویر و دسته‌بندی آنها استفاده گردد. در پژوهش حاضر، ضمن رسیدن به این هدف، نتیجه را در قالب یک پایگاه دانش و با استفاده از تعریف سه‌تایی‌ها ارائه نمودیم.

ج) یافتن مفهوم انتزاعی و بازنمایی آن به‌صورت روابط منطقی سبب می‌شود بتوان از این شیوه‌ی بازنمایی در جستجو و استنتاج معنایی استفاده کرد.

د) گاهی روابط معنایی انتزاعی میان تصاویر ممکن است بیش‌از یک مفهوم باشد. این شیوه‌ی تحلیل می‌تواند به یافتن روابط معنایی تصاویر کمک شایانی نماید.

منابع

- Bagheri, B.; Pourmohiabadi, M.; & Nezamabadipour, H. (1399). Content Based Image Retrieval by the Fusion of Short Term Learning Methods. *Iranian Journal of Electrical and Computer Engineering*, 4(13): 1-10. [In Persian]
- Baker, C.F.; Fillmore, C.J.; & Lowe, J.B. (1998) The Berkeley FrameNet project. In *Proceedings of the joint Annual Meeting of the Association for Computational Linguistics and International Conference on Computational Linguistics*, Montreal, QC, pp. 86-90.
- Barezi, Elham J.; & Kordjamshidi, P. (2024). Find the gap: Knowledge base reasoning for visual question answering. *arXiv*. <https://arxiv.org/abs/2404.10226>
- Barthes, R. (1968). *Elements of Semiology*. *Communications*, 4: 91-135, Hill and Wang: New York.
- Belcavello, F.; Timponi Torrent, T.; Matos, E. E.; Pagano, A. S.; Gamonal, M.; Sigiliano, N.; Dutra, L. V.; de Andrade Abreu, H.; Samagaio, M.; Carvalho, M.; Campos, F.; Azalim, G.; Mazzei, B.; de Oliveira, M. F.; Loçasso Luz, A. C.; Pádua Ruiz, L.; Bellei, J.; Pestana, A.; Costa, J.; Rabelo, I.; Silva, A. B.; Roza, R.; Souza, M.; & Oliveira, I. (2024). Frame2: A FrameNet-based multimodal dataset for tackling text-image interactions in video. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation*, pp: 7429–7437, Torino, Italia. ELRA and ICCL.
- Broadbent, D. E. (1953). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*, 47, 191–196.
- Chandler, D. (2022). *Semiotics: The Basics*. 4th ed. NY: Routledge.
- de Saussure, F. ([1916]1983). *Course in General Linguistics* (trans. Roy Harris). London: Duckworth.
- Diyanat, F. (1396). A study on the importance of semiotics (semantic significance) in conceptual photographs. *Theoretical Foundations of Visual Arts*, 4: 71-84. [In Persian]
- Emadoddin, A.R. (1396). Study of the Content Matching of Photos and News Content on the Websites of Al-Alam and Press TV News Networks. Master's Thesis. Faculty of Communication, University of Radio and Television of the Islamic Republic of Iran, Tehran. [In Persian]
- Emadoddin, A.R. (1399). Semiotics of news photography. *Journal of Rasaneh (Journal of Media Studies and Research)*, 31(1): 73-98. [In Persian]
- Fauconnier, G.; & Turner, M. (2002). *The Way We Think: Conceptual Blending*

- and the Mind's Hidden Complexities. New York: Basic Books.
- Fillmore, C.J. (1968). The case for case. In Emmon W. Bach and Robert T. Harms, editors, *Universals in Linguistic Theory*. Holt, Rinehart & Winston, New York, pp. 1-88.
- Fillmore, C.J. (1971) Some problems for case grammar. In R. J. O'Brien, editor, 22nd Annual Round Table. *Linguistics: Developments of the Sixties-Viewpoints of the Seventies*. Volume 24 of Monograph Series on Language and Linguistics. Georgetown University Press, Washington, D.C., pp. 35-56.
- Fillmore, C.J. (1982). Frame semantics. In *Linguistics in the Morning Calm*, Seoul, Korea: Hanshin, pp. 111-138.
- Fillmore, C.J. (1985) Frames and the semantics of understanding. In *Quaderni di Semantica*, 6.2:222-254
- Fillmore, C.J. (1994) Starting where the dictionaries stop: The challenge of corpus lexicography. In *Computational Approaches to the Lexicon*, ed. By B.T.S. Atkins and A. Zampolli, Oxford, pp. 349-393.
- Gildea, D.; & Jurafsky, D. (2002) 'Automatic labeling of semantic roles' In *Association for Computational Linguistics*, Vol. 28, Num. 3, pp245-288.
- Hakim, A.; Pakzad, Z.; & Kowsari, M. (1400). A study of verbal/visual multimodal discourse in contemporary Iranian art. *Quarterly Journal of Perspective*, 16 (60): 141-155. [In Persian]
- Halliday, M. A. K. (1994). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Hatefi, M.; & Shairi, H.R. (1390). The quasi-discursive status of the comparative semio semantics of text and image in the picture book of Ordinary People. *Journal of Comparative Art Studies*, 1 (2): 41-56. [In Persian]
- Hearst, M. (1999) 'Untangling text data mining' In *Proceedings of the 37th Annual Meeting of the ACL*, College Park, Maryland, pp. 3-10.
- Jewitt, C. (2009) *The Routledge Handbook of Multimodal Analysis*. London: RoutledgeFalmer.
- Kepes, G. (1403). *Images Language*. Translated by Firuzeh Mohajer. 18th edition. Tehran: Soroush Publications of the Islamic Republic of Iran Radio and Television. [In Persian]
- Kress, G.; Majdizadeh, Z.; & Hajjari, M. (2022). Multimodal discourse analysis. *Journal of Society, Culture, and Media*, 11(42): 279-310. [In Persian]
- Kress, G.; & Van Leeuwen, T. (2021). *Reading images: The grammar of visual design*. 3rd Eds. London: Routledge.
- Lin, T.Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., & Dollár, P. (2014) Microsoft COCO: Common Objects in Context. arXiv:1405.0312v3. <https://arxiv.org/abs/1405.0312>
- Mehdizadeh, A. (1392). Analysis of the act of photography and reading photos from a semiotic perspective. *The Glory of Art*, 3 (1): 1-74. [In Persian]
- Meunier, J.G. (2022). *Computational Semiotics*. London: Bloomsbury Publishing Plc.
- Padó, S. (2007). *Cross-lingual Annotation Projection Models for Role-Semantic Information*. PhD dissertation, Saarland University, Saarbücken, Germany.
- Peirce, C.S. (1931-58). *Collected Papers* (8 vols). Eds C. Hartshorne, P. Weiss & A. W. Burks. Cambridge: Harvard University Press.
- Pourghasem, H.; & Ghasemian, H. (1386). Semantic classification of medical images in a hierarchical structure based on a new unsupervised clustering method. In *Proceedings of the 13th Annual Conference of the Iranian Computer Association*. [In Persian]
- Razzaghi, P. (1397). Weakly supervised semantic segmentation using object level and context level information. *Journal of Machine Vision and Image Processing*, 5(1): 1-13. [In Persian]
- Ruppenhofer, J. and M. Ellsworth and M. R.L.Petruck and C. R.Johnson and J. Scheffczyk (2006) *FrameNet II: Extended Theory and Practice* <http://framenet.icsi.berkeley.edu/>

- Sadeghi, L. (1392). The blending of words and image in literary text based on conceptual blending theory. *Language Related Research*, 4 (3): 75-103. [In Persian]
- Shairi, H.R. (1397). *Visual Semiotics: Theories and Applications of Art Discourse*. Tehran: Sokhan Publications. [In Persian]
- Silverman, K. (1983). *The Subject of Semiotics*. New York: Oxford University Press.
- Sojoodi, F. (1387). *Applied Semiotics. Second Edition*. Tehran: Chavooshgaran Naghsh. [In Persian]
- Stevens, S. S. (1953). *Handbook of experimental psychology*. New York: John Wiley & Sons.
- Tanhayi Ahari, N.; Ameri, H.; Ameri, A.R.; & Shairi, H.R. (1399). The relationship between text and image in English Vision books based on discourse semantics. *Language Studies*, 12(37): 149-176. [In Persian]
- Xie, R.; Heinrich, S.; Liu, Z.; Weber, C.; Yao, Y.; Wermter, S.; Sun, M. (2020). Integrating image-based and knowledge-based representation learning. *IEEE Transactions on Cognitive and Developmental Systems*, 12(2): 169 – 178.
- Young, P.; Lai, A.; Hodosh, M.; & Hockenmaier, J. (2014). From image descriptions to visual denotations: New similarity metrics for semantic inference over event descriptions. *Transactions of the Association for Computational Linguistics*, 2:67–78.