****

**بسمه تعالی**

**خلاصه پيشنهاد سمينار کارشناسي ارشد**

**عنوان:پرسش و پاسخ تصویری**

**1- شرح مساله (با ارجاع به مراجع)**

در سال‌های اخیر پیشرفت‌های زیادی در مسائل هوش مصنوعی و یادگیری عمیق که در تقاطع دو حوزه پردازش زبان طبیعی و بینایی ماشین قرار می‌گیرند؛ رخ داده است. یکی از مسائلی که اخیراً مورد توجه قرارگرفته است. پرسش و پاسخ تصویری[[1]](#footnote-1) است. با توجه به یک تصویر و یک سؤال به زبان طبیعی، سیستم سعی می‌کند با استفاده از عناصر بصری تصویر و استنتاج جمع آوری شده از سوال متنی، پاسخ صحیح را پیدا کند. [1]

پرسش و پاسخ تصویری مرتبط با مسئله پرسش و پاسخ متنی[[2]](#footnote-2) است. پرسش و پاسخ متنی برای مدت طولانی در جامعه پردازش زبان طبیعی مورد مطالعه قرار گرفته است و پرسش و پاسخ تصویری نسخه گسترش یافته آن است که اطلاعات بصری به مسئله اضافه شده است. همین موضوع باعث ایجاد چالش هایی برای حل این مسئله شده است. تصاویر بعد بالاتر و نویز بیشتری نسبت به متن دارندعلاوه بر این، تصاویر فاقد ساختار و قواعد دستوری زبان هستند. در نهایت هم، تصاویر غنای بیشتری از دنیای واقعی را ضبط می‌کنند، در حالی که زبان طبیعی در حال حاضر نشانگر سطح بالاتری از انتزاع دنیای واقعی است. [2]

در این مسئله، ورودی یک تصویر و یک سؤال بر اساس تصویر است و خروجی یک یا چند کلمه است که به سؤال پاسخ می‌دهد.

**2- مباحث تحت پوشش سمينار(با ارجاع به مراجع)**

با توجه به موارد ذکر شده در قسمت شرح مساله، مباحث تحت پوشش این سمینار به ترتیب شامل موارد زیر خواهند بود:

1) بررسی کلی و تعاریف مبحث پرسش و پاسخ تصویری

2) کاربرد و اهمیت این مسئله

3) بررسی چالش‌های موجود در این مسئله

4) بررسی مجموعه دادگان مطرح و مسابقات مطرح این حوزه

5) ارزیابی مدل‌های ارائه شده

5) چگونگی ساخت مجموعه داده حاوی پرسش و پاسخ به زبان فارسی

**3- اهميت موضوع**

کاربردهای بی‌شماری برای VQA وجود دارد. یکی از مهم‌ترین مواردی که در بین آن‌ها وجود دارد این است که از VQA می‌توان به عنوان دستیار برای افراد کم‌بینا و نابینا استفاده کرد.

**4**- نتيجه ارزيابي در گروه:

قبول رد تصحيح ارسال براي داوري تاريخ --/--/---- امضاء مدير گروه:

# مراجع

|  |  |
| --- | --- |
| [1] | Yash Srivastava and Vaishnav Murali and Shiv Ram Dubey and Snehasis Mukherjee, "Visual Question Answering using Deep Learning: A Survey and Performance Analysis," *arXiv,* vol. abs/1909.01860, 2019. |
| [2] | Qi Wu and Damien Teney and Peng Wang and Chunhua Shen and Anthony R. Dick and Anton van den Hengel, "Visual question answering: A survey of methods and datasets," *ArXiv,* vol. abs/1607.05910, 2017. |

سناریو

نوع دیتا

کاربرد ها اهمیت

های متفاوت vqa flavor

چطور یک ایجنت گفت و گو میتونه که واقعی تر بشن ؟ با استفاده از تصویر

ربات ها برای انجام و امر و نهی

1. Visual Question Answering [↑](#footnote-ref-1)
2. Textual Question Answeing [↑](#footnote-ref-2)