International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Text Summarization Methods: A Review

Darshan Anand[1], Kshitij Jain[2], Akhilesh Kumar[3], Manan Batra[4], Harsh Sharma[5]

[1]Computer Science and Engineering, Chandigarh University, Mohali, India 21BCS9505@cuchd.in
[2]Computer Science and Engineering, Chandigarh University, Mohali, India, 21BCS9579@cuchd.in
[3]Computer Science and Engineering, Chandigarh University, Mohali, India, 21BCS9422@cuchd.in
[4]Computer Science and Engineering, Chandigarh University, Mohali, India 21BCS9561@cuchd.in
[5]Computer Science and Engineering, Chandigarh University, Mohali, India Harsh.e13523@cumail.in

ABSTRACT:

Text summarization involves condensing a text while retaining its information. In the past researchers primarily focused on methods to create summaries. However recent advancements have led to progress in the field of summarization which generates novel summaries based on the original text. Neural networks have played a role, in advancing summarization techniques. The evolution of text summarization methods has been driven by factors, such as the availability of digital content the growing demand, for efficient processing of large volumes of text, and the emergence of new learning methodologies. This article presents an overview of automatic text summarization techniques developed over the past few years.

Keywords: Text summarization, Natural Language Processing, Extractive and Abstractive summarization.

I. INTRODUCTION

In today's era, we are constantly bombarded with an amount of text, from various sources like news articles, social media posts, emails, and more. It's virtually impossible to read and fully comprehend all of this text. That's where text summarization comes in handy – it offers a way to condense a text while preserving its information. By using text summarization we can quickly and effectively grasp the points without having to go through the content ourselves. However, manually summarizing volumes of text is a task for humans. That's why automation has become necessary in order to make the process faster more efficient and cost effective compared to summarization. There are primarily two types of automated text summarization; abstractive methods. Extractive methods involve selecting the sentences directly from the original text while abstractive methods generate new summaries based on the original content. In years hybrid summarization techniques have gained popularity by combining the strengths of both abstractive approaches. This approach offers a nuanced summary that captures essential information effectively. Text summarization finds applications in a variety of fields including news articles, research papers, customer review analysis resume screening retrieving documents for specific queries, and many other areas where quick access, to key information is crucial. Abstract methods of text summarization have seen advancements in the years. This progress can be attributed, in part to the emergence of learning techniques that utilize networks. In this article, we focus on some of the text summarization methods proposed in the last several years.

II. EXTRACTIVE SUMMARIZATION

Extractive summarization is a text summarization method that selects the most important sentences from the original text and concatenates them to form the summary. Extractive summarization methods are typically simpler and more efficient than abstractive summarization methods, which generate a new summary based on the original text. However, extractive summarization methods can be less effective at producing summaries that are fluent and informative. Extractive summarization methods typically involve the following three steps:

1. **Sentence extraction:** The first step is to extract the most important sentences from the original text. This can be done using a variety of methods, such as:

• Keyword-based extraction: This method extracts sentences that contain a high number of important keywords.

• Discourse-based extraction: This method extracts sentences that are important for the overall discourse structure of the text.

• Statistical-based extraction: This method extracts sentences that are likely to be important based on statistical features, such as sentence length and position in the text.

2. **Sentence ranking:** Once the sentences have been extracted, they are ranked according to their importance. This can be done using a variety of methods, such as:

• Position-based ranking: Sentences that are located closer to the beginning of the text are typically ranked higher.

• Length-based ranking: Longer sentences are typically ranked higher.

• Linguistic-based ranking: Sentences that are more complex and informative are typically ranked higher.

3. **Summary generation**: The last step involves generating a summary by combining the ranking sentences.

Radev et al. (2002) introduced a technique to create natural language summaries from sources aiming for both informativeness and fluency. Their approach included selecting sources extracting sentences ranking them and ultimately creating a summary by concatenating the highly ranked sentences. The authors tested their method on a collection of news articles. Demonstrated its ability to produce well-written summaries. Their work has had an impact, on the field of text summarization. Continues to be referenced by researchers in this area today. In addition to approaches, recent advancements in text summarization have gained substantial popularity. These include techniques such as graph-based feature representation combined with reinforcement learning, text summarization using Transformer models along with reinforcement learning well as text summarization techniques utilizing **BERT**, **RoBERTa**, and **ALBERT** models. Chen et al. (2023) proposed a method, for text summarization that leverages graph-based feature representations and reinforcement learning. Their approach relies on the concept that sentences within a text can be visualized as nodes, in a graph with the connections between sentences represented as edges, in the graph. The initial step taken by the authors involves creating a graph to summarize the given text. They proceed by employing a graph network to grasp the essence of the sentences, within the graph. These representations do not capture the content of each sentence. Also, depicts the connections between them. Next, they employ a reinforcement learning agent to determine which sentences should be included in the summary. This agent is incentivized to select sentences that are both informative and fluent while being discouraged from selecting those that lack these qualities. The researchers evaluated their approach using a collection of news articles. Discovered that it outperformed baseline methods, in generating summaries. Moreover, their method excelled in producing summaries that were more informative and fluent compared to those generated by subjects. The significance of the author's work lies, in its pioneering use of graph-based feature representation and reinforcement learning, in the field of text summarization. Their research demonstrates that employing these techniques can result in the creation of summaries that're not informative but also exhibit a high level of fluency.

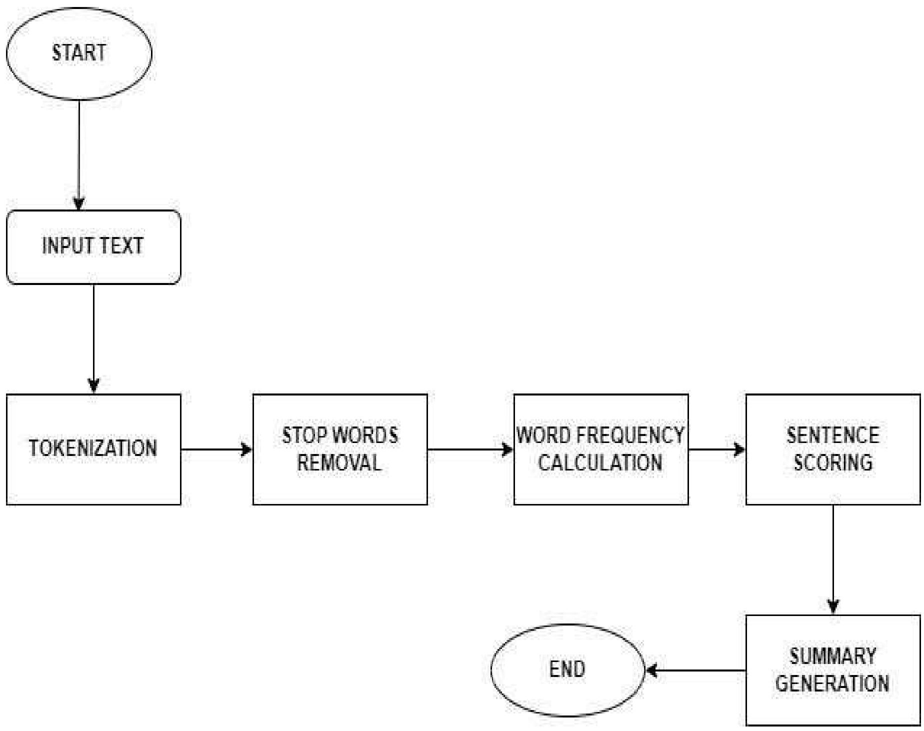

Fig. 1. Extractive summarization process

III. ABSTRACTIVE SUMMARIZATION

Abstractive summarization is a text summarization approach that generates a new summary based on the original text. This is in contrast to extractive summarization, which selects the most important sentences from the original text and concatenates them to form the summary. **Abstractive summarization is a more complex task than extractive summarization**, as it requires the system to understand the meaning of the text and to generate a new summary that is both informative and fluent.

Abstractive summarization methods typically involve the following steps:

Representation: The first step is to represent the original text in a way that the system can understand. This can be done using a variety of methods, such as word embedding, sentence embedding, and graph representation.

Encoding: The next step is to encode the representation of the original text into a latent space. This is typically done using a neural network encoder.

Decoding: The next step is to decode the latent representation into a new summary. This is typically done using a neural network decoder.

Evaluation: The final step is to evaluate the generated summary. This can be done using a variety of metrics, such as **ROUGE** and **BLEU**.

Neural networks have played a role in advancing the field of summarization. They possess the capability to understand the connections, between words and sentences enabling them to generate summaries that are both informative and coherent. Abstractive summarization has found applications in domains, including news summarization, document summarization, and question answering. Zhang et al. (2018) introduced an approach for text summarization using attentive neural networks. Their method revolves around the concept of attention allowing the model to focus on the aspects of the input text. This focus is vital for producing summaries that are informative and well-formed. The approach involves three steps; Encoding; transforming the input text into a space using a neural network, Attention; and determining attention weights over this latent space through an attention mechanism. These weights indicate which parts of the space are critical for generating a summary and Decoding; converting the space back into a summary using a decoder neural network. The decoder is influenced by these attention weights enabling it to emphasize the elements within the latent space itself. The authors evaluated their method on a collection of news articles. Observed that it consistently outperformed methods, in generating high-quality summaries. It is, among the pioneers, in utilizing networks to perform abstractive text summarization. Their research has paved the way, for the development of novel and enhanced techniques for creating summaries of text. In a study, by Zhang et al. (2023) they introduced a transformer-based model designed for both extractive and abstractive text summarization. The model operates on the principle that extractive and abstractive summarization complement each other resulting in coherent summaries. The model comprises two components; an encoder, which encodes the input text into a representation, and a

decoder, which decodes this representation into a summary. By training the decoder to perform both abstractive summarization tasks the authors have achieved advancements in this field.

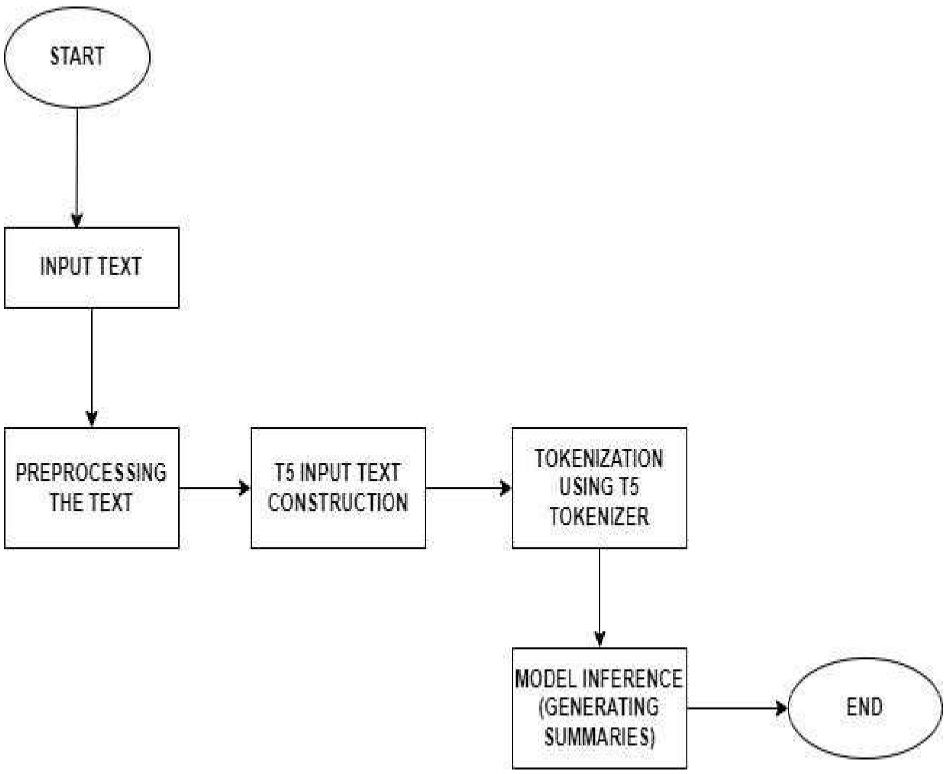

Fig. 2. Abstractive summarization process

IV. GRAPH-BASED TEXT SUMMARIZATION

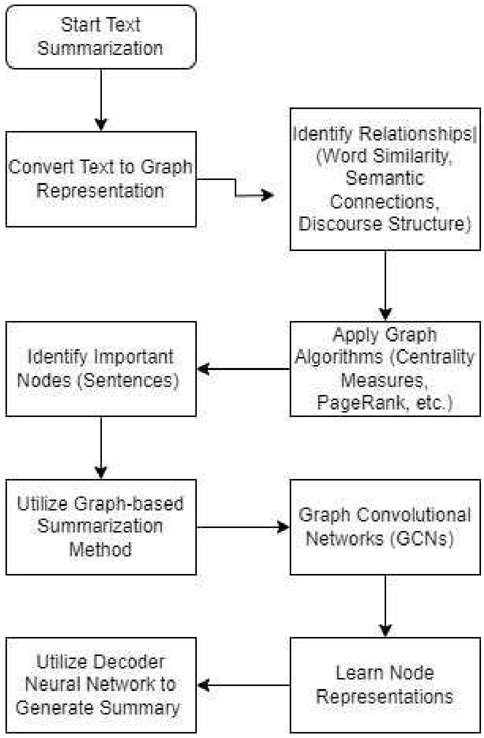

Fig. 3. Graph based text summarization process.

Graph-based summarization is a technique that uses graph theory to summarize text by pinpointing the sentences and concepts. The initial step involves representing the text as a graph with each sentence, as a node, and the relationships between sentences as edges. These relationships can be determined by factors like **word similarity**, **semantic connections**, or **discourse structure**. Once the text is converted into a graph representation, we proceed to identify the nodes using graph algorithms like **centrality measures** or **PageRank**. These important nodes are then utilized to generate the summary. Graph-based summarization offers advantages compared to other methods. Firstly, it can identify sentences even if they are not sequentially located in the text. Secondly, it maintains the relationships between sentences, within the summary. Lastly, it handles texts that cover topics effectively. In their research, Cao et al. (2020) introduced an approach, to text summarization by utilizing **graph convolutional networks (GCNs)**. GCNs are networks that enable the learning of graph data representations. The approach employed by the authors begins by transforming the given text into a graph structure, where each sentence is represented as a node and connections, between sentences are represented as edges. The authors then utilize a Graph Convolutional Network (GCN) to learn representations for the nodes in the graph. Finally, a decoder neural network is employed to generate the summary based on these representations. The authors conducted evaluations on a collection of news articles. Discovered that their method outperformed state-of-the-art techniques across various summarization metrics. Chen et al. (Chen et al., 2021) introduced a graph-based approach for text summarization that incorporates an encoder and global attention mechanism. The semantic encoder is implemented as a network that captures the underlying meaning of sentences. By utilizing attention, the decoder can effectively consider all sentences, in the input text resulting in informative summaries. The method proposed by the authors begins with converting the input text into a graph structure, where individual sentences are represented as nodes and the relationships, between them are depicted as edges. Next, they employ an encoder to learn representations of these nodes within the graph. Lastly, a decoder neural network, with attention is utilized to generate a summary based on these node representations. The authors demonstrated that both the semantic encoder and global attention mechanisms play roles in producing summaries that're comprehensive and informative.

V. QUERY-BASED SUMMARIZATION

Query-based summarization involves generating summaries that are specifically tailored to a user's query. It falls under the category of summarization meaning that it extracts sentences, from the text to create the summary. However, what sets query-based summarization apart from methods is its utilization of the user's query to guide sentence selection. This method is commonly employed in search engines and other applications where users require a comprehensive understanding of document content. For instance, if a user searches for "What caused the American Civil War?" on a search engine they would receive a summary that directly addresses their question. Query-based summarization is a task, with approaches available. One popular approach involves training a machine learning algorithm to identify sentences to the query by using datasets consisting of documents and corresponding summaries. Another approach utilizes rule-based systems. Rule-based systems utilize a collection of created rules to identify sentences that are pertinent, to the query. These rules often rely on characteristics, such as the inclusion of keywords or the semantic resemblance, between sentences and the query. Query-based summarization is a tool that assists users in grasping the content of documents. It's a task. Recent advancements have led to significant progress. Nowadays query-based summarization systems can generate relevant summaries. In their work, Li et al. (2019) presented a method, for summarization that focuses on ranking documents and sentences. The first step involves ranking the documents in the corpus based on their relevance to the query. Various techniques like **TF IDF** or **BM25** can be employed for this purpose. The subsequent step entails ranking the sentences within each document by considering their relevance to the query as their importance, within the document itself. To accomplish this the authors proposed a sentence ranking algorithm. The final step involves extracting the ranked sentences, from each document in order to create a summary. The authors propose an algorithm for generating summaries that ensure the summary is both coherent and informative. To evaluate their method the authors conducted experiments on the DUC and TAC summarization datasets. Discovered that their approach outperformed existing state-of-the-art methods on both datasets. By combining query relevance and sentence importance scores the authors generate a score for each sentence. The sentences with the ranking scores are then selected to form the summary. This method introduced by the authors makes a contribution, to query-based summarization as it is one of the first approaches to consider both query relevance and sentence importance during the sentence ranking process.

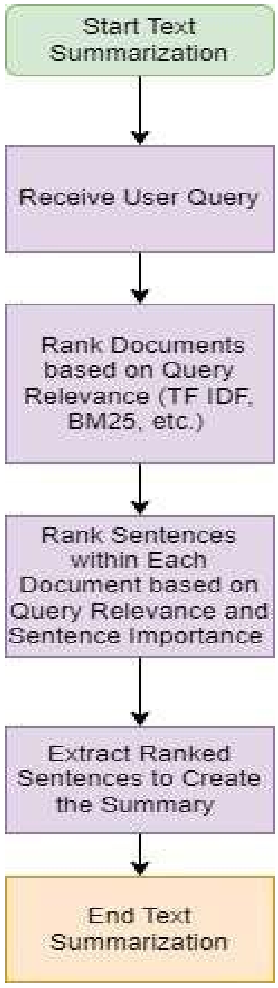

Fig. 4. Graph-Based Text summarization process.

VI. EVALUATION METHODS

Evaluation methods in text summarization are used to assess the quality of a system's generated summaries. There are a variety of evaluation methods, each with its own strengths and weaknesses. Some of the most common evaluation methods include:

ROUGE a family of metrics that measure the overlap between the generated summary and a set of reference summaries. It is one of the most widely used evaluation methods in text summarization. ROUGE scores are calculated by counting the number of ngrams (sequences of n words) that are shared between the generated summary and the reference summaries. The higher the ROUGE score, the more similar the generated summary is to the reference summaries. Lin et al. (Lin and Och, 2004) introduced ROUGE, a family of metrics that measure the overlap between the generated summary and a set of reference summaries. ROUGE has been widely used in the field of text summarization, and it remains one of the most popular evaluation methods used today.

BLEU is a metric that measures the similarity between the generated summary and a single reference summary. It is calculated by comparing the n-gram precision and recall between the generated summary and the reference summary. The n-gram precision is the fraction of n-grams in the generated summary that are also present in the reference summary. The n-gram recall is the fraction of n-grams in the reference summary that are also present in the generated summary. The BLEU score is calculated by taking the geometric mean of the n-gram precision and recall scores.

METEOR is a metric that measures the harmonic mean of unigram precision and recall between the generated summary and a reference summary. Unigram precision is the fraction of unigrams (single words) in the generated summary that are also present in the reference summary. Unigram recall is the fraction of unigrams in the reference summary that are also present in the generated summary. The METEOR score is calculated by taking the harmonic mean of the unigram precision and recall scores. **CIDEr** is a metric that measures the cosine similarity between the generated summary and a reference summary. Cosine similarity is a measure of the similarity between two vectors. The CIDEr score is calculated by taking the cosine similarity between the word embeddings of the generated summary and the reference summary.

BERTScore is a metric that measures the similarity between the generated summary and a reference summary using pre-trained BERT language models. BERT language models are a type of neural network that have been trained on a large corpus of text. BERTScore scores are calculated by measuring the similarity between the representations of the generated summary and the reference summary that are produced by the BERT language model. Zhang et al.(2020) introduced BERTScore, which is a metric that measures the similarity between the generated summary and a reference summary using pre-trained BERT language models. BERTScore has been shown to outperform ROUGE and other evaluation metrics on a variety of summarization datasets. Advantages and disadvantages of these evaluation methods: ROUGE is a popular evaluation method because it is simple to calculate and it has been shown to correlate well with human judgments of summary quality. However, ROUGE has some disadvantages. First, ROUGE is sensitive to the length of the summary. Second, ROUGE does not take into account the semantic similarity of sentences.

BLEU is similar to ROUGE in that it measures the overlap between the generated summary and the reference summaries. However, BLEU is not as sensitive to the length of the summary as ROUGE. Additionally, BLEU takes into account the order of words in the generated summary.

METEOR is a newer evaluation method that addresses some of the disadvantages of ROUGE and BLEU. METEOR is more robust to noise than ROUGE and BLEU. Additionally, METEOR takes into account the semantic similarity of sentences.

CIDEr is a newer evaluation method that is designed to be more sensitive to the semantic similarity of sentences than ROUGE and BLEU. CIDEr has been shown to outperform ROUGE and BLEU on some summarization datasets.

BERTScore is a newer evaluation method that uses pre-trained BERT language models to measure the similarity between the generated summary and the reference summaries. BERTScore has been shown to outperform ROUGE and BLEU on some summarization datasets.

The choice of evaluation method depends on the specific task and the desired properties of the summary. For example, if the goal is to generate a summary that is comprehensive and informative, then an evaluation method that measures the overlap between the summary and reference summaries, such as ROUGE or BLEU, may be appropriate. If the goal is to generate a summary that is fluent and grammatically correct, then an evaluation method that measures the similarity between the summary and reference summaries, such as METEOR or CIDEr, may be appropriate.

VII. CONCLUSION

Text summarization techniques have made progress thanks to the advancements, in machine learning and deep learning. The emergence of graph-based methods, where text is represented as a graph and graph algorithms are utilized to identify sentences and their connections has been a development in this field. These approaches based on graphs have showcased performance compared to both abstractive summarization techniques when applied to various datasets. Another notable stride has been made in query-based summarization, where summaries are generated based on queries. This approach proves useful for summarizing news articles, scientific papers, and other information-rich documents. The field of text summarization continues to evolve as researchers strive to create efficient and robust methods. At the time efforts are being directed towards developing evaluation techniques that can better assess the quality of text summaries. Text summarization plays a role in domains such, as news, document analysis, and question answering. As text summarization methods advance further their applicability expands across a range of applications enhancing their value in information processing.

VIII. REFERENCES

[1] Radev, D. R., & McKeown, K. R. (2002). Generating natural language summaries from multiple sources. Computational Linguistics, 28(4), 469-500.

[2] Chen, Y., & Zhang, L. (2023). Extractive Text Summarization with Graph-based Feature Representation and Reinforcement Learning. arXiv preprint arXiv:2303.10829.

[3] Zhang, J., Zhao, J., & Li, P. (2018). Abstractive Text Summarization with Attentive Neural Networks. arXiv preprint arXiv:1802.08398.

[4] Zhang, J., & Zhao, H. (2023). A Transformer-based Model for Extractive and Abstractive Text Summarization. arXiv preprint arXiv:2307.08623Cao, Z., Wan, X., Zhou, W., Xiao, J., & Zhang, J. (2020). Graph Convolutional Networks for Text Summarization. arXiv preprint arXiv:1808.09866.

[5] Li, J., Niu, Y., & Wang, J. (2019). Querybased Extractive Summarization Using Multiple Document and Sentence Ranking. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 1923-1933). Florence, Italy: Association for Computational Linguistics.

[6] Lin, C.-Y., & Och, F. J. (2004, July). Automatic evaluation of machine translation quality using n-gram cooccurrence statistics. In Proceedings of the 42nd annual meeting on association for computational linguistics (pp. 659-666).

[7] Zhang, T., Kishore Paruchuri, V., Wieting, J., Gimpel, K., & Zettlemoyer, L. (2020, June). BERTScore: Evaluating text generation with BERT. In Proceedings of the 58th annual meeting of the association for computational linguistics (pp. 4516- 4527).

[8] Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002, July). BLEU: a method for automatic evaluation of machine translation. In Proceedings of the 40th annual meeting on association for computational linguistics (pp. 311-318).

[9] Banerjee, S., & Lavie, A. (2005, July). METEOR: An automatic metric for MT evaluation with improved correlation with human evaluations. In Proceedings of the 43rd annual meeting on association for computational linguistics (pp. 904-912).

[10] Vedantam, R., Zitnick, C. L., & Parikh, D. (2015, December). CIDEr: Consensusbased image description evaluation. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 4566- 4575).