Project Literature Review Group-45

Guess it! An Online Multiplayer Trivia

Varun Musunuru
AM.EN.U4CSE19336

Sai Pavan Krishna Nagam
AM.EN.U4CSE19347

Venkata Praneeth Villuri
AM.EN.U4CSE19358

Gopikrishna Vasantha
AM.EN.U4CSE19359

Mentor: Dr. Thushara MG

Amrita Vishwa Vidyapeetham, Amritapuri, Kerala, India
B.Tech, Computer Science and Engineering

1 Introduction

Traditional approaches for searching for a keyword have largely been replaced by keyword extractors thanks to the rapid development in NLP community. Now, searching for an appropriate keyword extractor has become a hassle as each of them has its own way of working. We looked into several keyword extractors like KeyBERT, RAKE, YAKE,sentence transformers, and TextRank(summa). To gain a complete image understanding, we should not only concentrate on classifying different images but also try to precisely estimate the concepts and locations of objects contained in each image. This task is referred as object detection. As one of the fundamental computer vision problems, object detection is able to provide valuable information for semantic understanding of images and videos.

2 Papers

2.1 YAKE! Collection-Independent Automatic Keyword Extractor

In this paper, they talked about YAKE, one of the widely used keyword extractor since YAKE supports different domains and languages. Unlike others, it does not rely on dictionaries. They talked about how it uses an unsupervised approach to build upon features directly from the given text, which makes it a whole lot of flexibility compared to other keyword extractors. So, it is possible to use it in several languages. They even compared YAKE to other keyword extractors like RAKE and IBM NLU (IBM Natural Language Understanding).

Even though rapid progress has been made in extracting keywords, it is yet to be properly solved and this YAKE is one step towards it. As many keyword extractors follow the traditional supervised learning model which requires training, YAKE can be instantly applied to documents across several languages. They also mentioned that the ongoing demand for YAKE is because of their plug-and-play nature since they can be applied easily to several documents.

They talked about the keyword extraction pipeline for YAKE which consists of six main components namely:

1. Text pre-processing
2. Feature Extraction
3. Individual terms score
4. Candidates keyword list generation
5. Data De-duplication

6. Ranking

In the preprocessing part, the text is split into individual terms based on the delimiter(e.g., linebreaks, brackets, comma, period, etc). Next, a set of five features has been created to capture the characteristics of each term. These are:

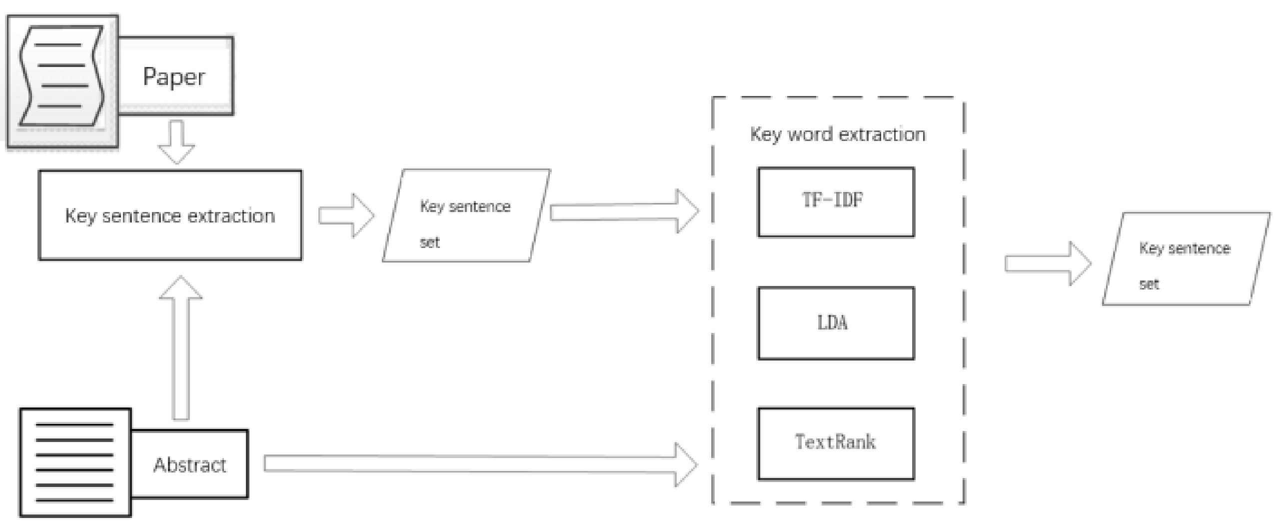
- Casing;
- Word Positional;
- Word Frequency;
- Word Relatedness to Context; and
- Word DifSentence.

Casing reflects the casing of a word. Word positional values are calculated based on the fact that important words occur at the beginning of the sentence. word frequency indicates the frequency of the word. Word relatedness to context computes the number of different terms that occur to the side of the candidate word. Finally, the word DifSentence signifies how often a candidate word occurs in the text.

In the third step, we heuristically combine all these features into a single measure such that each term is assigned a score S(w). This weight is taken into consideration to generate the candidate words. Here, we consider a sliding window of 3-grams, thus generating a contiguous sequence of 1, 2, and 3-gram candidate keywords. Each candidate keyword will then be assigned a final S kw , such that the smaller the score the more meaningful the keyword will be.

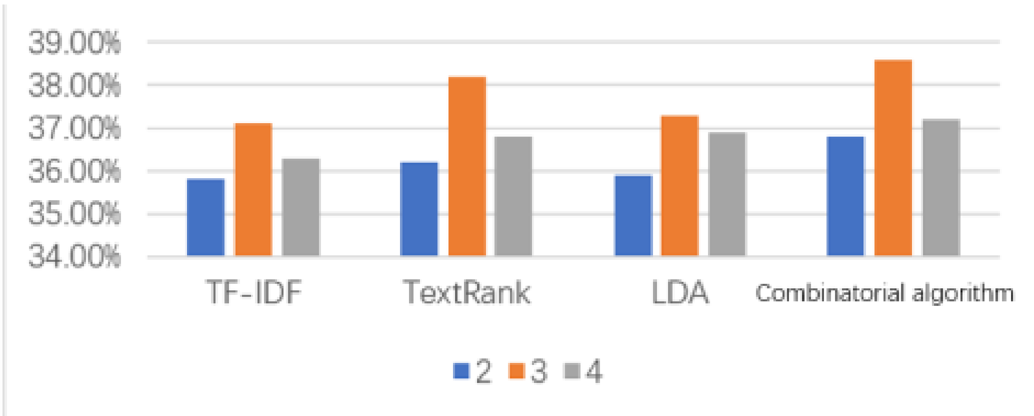
2.2 Bert-Based Text Keyword Extraction

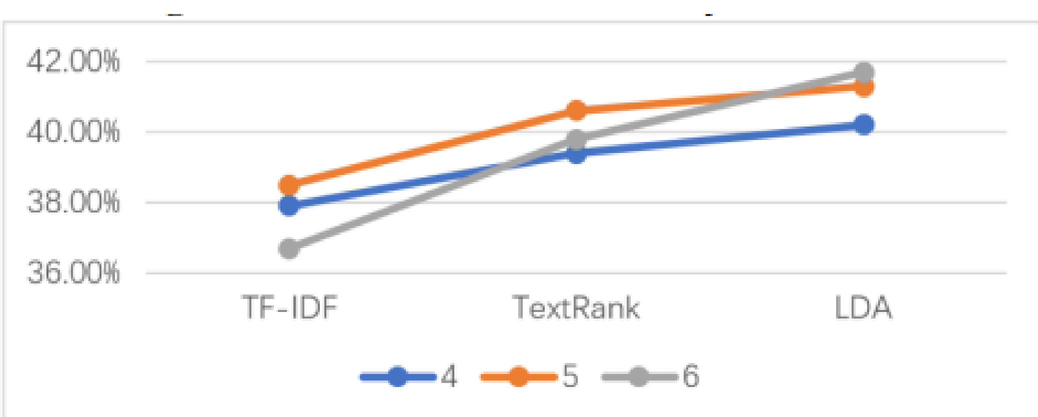
In recent years, text keyword extraction has been widely used in various fields of Natural language Processing like text topic mining, text classification, dialogue system generation, text similarity calculation, and many more. This paper proposed a method for extracting keywords based on Bert and Multiclass feature fusion. Generally, if we want to give a description of a paper the keywords extracted are less due to the small abstract. In this method, Using the Bert model first they extracted the sentences with high semantic similarity to the abstract from the paper which is used as the background material to expand the abstract. Then the first K keywords are extracted from the new abstract using TF-IDF, Text rank, and LDA algorithms based on multi-category features. The final keyword set is obtained by the idea of intersection.



BERT (Bidirectional Encoder Representations from Transformers) makes use of Transformer, an attention mechanism that learns contextual relations between words in a text. This model can consider the full context of a word by looking at the words that come before and after it, which is particularly useful for understanding the intent behind search queries. First, document embedding (a representation) is generated using the sentences-BERT model. Next, the embedding of words is extracted for N-gram phrases. The similarity of each keyphrase to the document is then measured using cosine similarity. The most similar words can then be identified as the words that best describe the entire document and are considered keywords.

For key sentence extraction, for each sentence in the abstract the Bert modal is used to find n sentences with the highest semantic similarity. Then all the sentences are de-duplicated and sorted and the top N sentences are extracted as the final key sentences. Using these Final Key sentences and abstract the keywords are extracted using the three algorithms. These keywords are unionized and de-duplicated to form the final keyword set.





This is a graph representation of performance by various algorithms for different numbers of key sentences and different numbers of keywords to select optimal values.

	Evaluating indicator	TF-IDF	Text rank	LDA	The algorithm in this paper
Key sentences are not extracted	P	37.3%	39.2%	39.8%	40.1%
	R	39.1%	42.2%	43.0%	44.2%
	F	38.2%	40.6%	41.3%	42.1%
Extract key sentences	P	36.9%	38.6%	40.6%	*42.3%*
	R	40.1%	43.0%	43.5%	*45.0%*
	F	36.4%	40.7%	42.0%	*43.6%*

Here we can see the difference in recall rate, accuracy rate, and f values in both cases. there is a significant raise in the values when we use the Bert algorithm.

2.3 A Comparative Study on different Keyword Extraction Algorithms

A Comparative Study on different Keyword Extraction Algorithms

The growth in research paper publications is pretty evident. So, searching relevant documents related to our domain and finding the needed topic from the paper has become a tedious task. That's where the keywords and keyphrases come into play this paper dwells on different keyword extractors and a comparative study has been done. It compares the performance of PositionRank which considers the position of all words occurrences in the document with TextRank and RAKE (Rapid Automatic Keyword Extraction).

Machine Learning approaches for keyword extractions have been discussed, namely

- Supervised
- Unsupervised
- Semi-Supervised

A brief insight about the algorithms studies has been given. TextRank algorithm scores keywords based on the co-occurrence connections between words. In Rapid Automatic Keyword Extraction (RAKE), keywords have multiple content words which are informative rather than punctuation and stop words and lastly, in position rank which is an unsupervised model, keywords and keyphrases are extracted from documents that contain information from all positions of the occurrences of a word.

The use of TF*IDF has been discussed in supervised learning.TF*IDF scores the keyphrases based on their rate of occurrence in the research paper and the number of times the word or a phrase occurs in the corpus of the same kind of research paper. The words which are having less TF*IDF value are significant. This cant be the same for the unsupervised approach as only term frequency is considered because when a single document is considered the IDF score would be the same for all candidate words.

In Rapid Automatic Keyword Extraction (RAKE), keywords have multiple content words which are informative rather than punctuation and stop words. A detailed explanation of the working of RAKE has been discussed meticulously.

- Candidate Selection
- Feature Calculation
- Keyword Selection

Position Rank, an unsupervised graph-based algorithm, considers both the position of the word and its frequency in a document for keyphrase extraction. All positional occurrences of words are considered in position rank which performs better than the methods which consider the only first position of a word.

Coming to Text Rank, this is a graph-based model which uses the PageRank algorithm. It is used to extract the summary of a text. After preprocessing the text, a vocabulary of unique words is formed where each word acts as a vertex. A weighted undirected edges graph is built. Information about the connections(edges) between all vertices will be present in the weighted edge matrix.

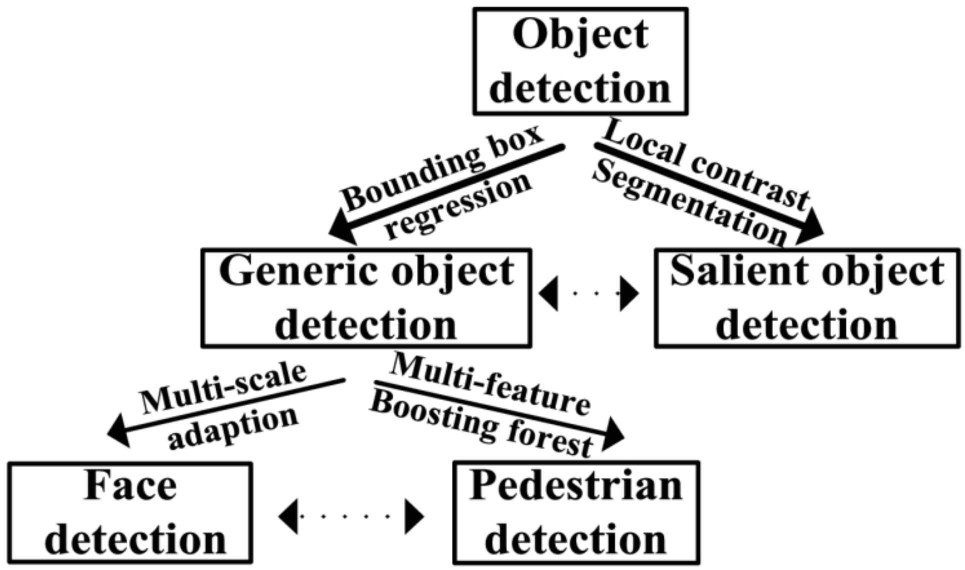
Coming to the results, they have experimented with the mentioned algorithms with large size research documents for determining key phrases. They inferred that certain locations like title, and abstract are where keywords appear mostly. They came to the conclusion that PositionRank gives more weight to the words which occur in the initial positions of a document similar to the TextRank Algorithm whereas In the RAKE algorithm, the highest scoring phrases which they got as output are quite long, less relevant and the score of relevant phrases is low.

Thanks to the insights provided we came to the conclusion that Position rank considers both position and frequency of a word so it gives better results when compared to both TextRank and RAKE.

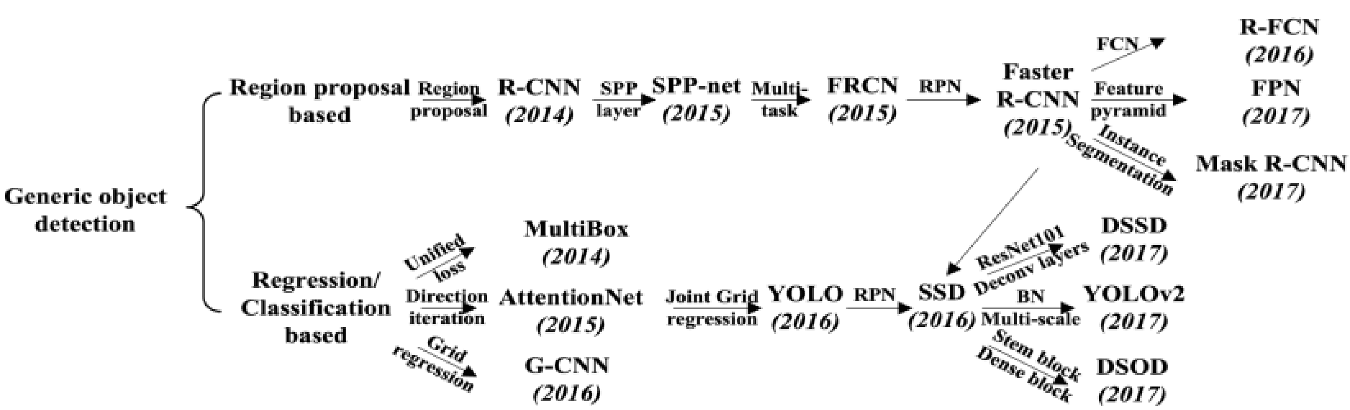
2.4 Object Detection With Deep Learning: A Review

Due to its powerful learning ability and advantages in dealing with occlusion, scale transformation, and background switches, deep learning-based object detection has been a research hotspot in recent years. This paper provides a detailed review on deep learning-based object detection frameworks that handle different subproblems, such as occlusion, clutter, and low resolution, with different degrees of modifications on R-CNN. The review starts on generic object detection pipelines which provide base architectures for other related tasks. Then, three other common tasks, namely, salient object detection, face detection, and pedestrian detection, are also briefly reviewed.

Generic object detection aims at locating and classifying existing objects in any one image and labeling them with



rectangular Bounding Boxes to show their confidence in existence. The frameworks of generic object detection methods can mainly be categorized into two types. One follows the traditional object detection pipeline, generating region proposals at first and then classifying each proposal into different object categories. The other regards object detection as a regression or classification problem, adopting a unified framework to achieve final results (categories and locations) directly. The region proposal-based methods mainly include R-CNN, spatial pyramid pooling (SPP)-net, Fast R-CNN, Faster R-CNN, region-based fully convolutional network (R-FCN), feature pyramid networks (FPN), and Mask R-CNN, some of which are correlated with each other (e.g., SPP-net modifies R-CNN with an SPP layer). The regression/classification-based methods mainly include MultiBox, AttentionNet, G-CNN, YOLO, Single Shot MultiBox Detector (SSD), YOLOv2, deconvolutional single shot detector (DSSD), and deeply supervised object detectors (DSOD). The correlations between these two pipelines are bridged by the anchors introduced in Faster R-CNN. Details of these methods are as follows.



3 Summary

3.1 Comparison on Keyword Extraction Algorithms

Algorithm	Description	Advantages
Rake	Domain-Independent keyword extraction algorithm	Easily applicable to new fields, Very effective in dealing with multiple types of documents, Faster but less accurate than YAKE and KeyBert.
Yake	Lightweight, unsupervised automatic keyword extraction method	Easy to implement, Supports different domains and languages, Do not require training on external resources.
KeyBert	Find sub-phrases using BERT-embeddings and simple cosine similarity	Not faster as RAKE but more accurate, With addition of vectorization gives sensical keywords, Do not require training on external resources.
Text Rank	Unsupervised graph-based approach	Easy to implement, Faster and lighter, Do not require training on external resources.

3.2 Comparison on Object Detection Algorithms

Algorithm	**Authors**	**Year**	**Approaches**	**Advantages**	**Limitations**
R-CNN	R. Girshick et al.	2014	Region Proposal Generation. CNN-Based Deep Feature Extraction.	Improved the quality of candidate BBs. Extracted high-level features.	Training is expensive in space and time. The obtained region proposals are still redundant.
SPP-net	K. He et al	2015	Takes several finer to coarser scales to partition the image. Reuses feature maps of the 5th conv layer.	Improved detection efficiency. Better results with correct estimations.	Multi-stage pipeline. Additional expense on storage space.
FRCN	R. Girshick	2015	Multi-task loss on classification. Bounding box regression	Processed in single stage. Saves the additional expense of storage.	Rely on additional methods to generate regions. Model efficiency depends on region proposal.
Faster R-CNN	S. Ren et al.	2015	Introduced Region Proposal Network.	Nearly cost-free way by sharing full-image conv features with detection network.	The alternate algorithm is time consuming. Model is not skilled in dealing with extreme scales.
R-FCN	J. Dai et al.	2016	Produced position-sensitive score maps. Class-agnostic Bounding boxes.	Is fully convolutional with almost all computation shared on the entire image.	Training time and space increases rapidly while improving scale invariance.
FPN	T.-Y. Lin et al.	2017	Builds image feature pyramids in Bottom-up, Top-down and several other ways at different depths.	Extracts rich semantics from all levels. Trained end to end with all scales.	The locations of these features are not precise, because these maps have been downsampled and upsampled several times.
Mask R-CNN	K. He et al.	2017	Predicts segmentation masks in a pixel to pixel manner. Adopts a simple and quantization-free layer.	Instance segmentation is achieved. Preserve the explicit per-pixel spatial correspondence faithfully.	The mask branch only adds a small computational burden and its cooperation with other tasks provides complementary information for object detection.
YOLO	J. Redmon et al.	2016	Uses top-most feature map to predict both confidences and Bounding Boxes. Object detection task is transformed into regression problem.	Faster in detection. Real-time detectors and Recognitions. Fewer false positives on the background.	The position is not accurate, and the detection effect of small and dense instances is not efficient. Difficulty in generalize to objects in new/unusual aspect ratios.
SSD	W. Liu et al.	2016	Multi-scale detection is realised by using feature layers from different scales.	Eliminates proposal generation. Faster while providing unified framework for both training and inference.	Due to the deep convolutional layer, the extracted features may be lost for smaller targets.

4 Conclusion

As part of the game engine, we need to extract keywords from the description of the movie/series which can be used to guess the movie. There are many algorithms that were proposed in this field of keyword extraction, out of the Rake, Yake, KeyBert, and Text Rank algorithms were tested against a few sample movie descriptions. Out of these four Rake and KeyBert extracted meaningful keywords relevant to our application.

The next step in the development of our game engine deals with object detection. Object Detection algorithms based on deep learning have been widely applied in many fields. We also employ this in our game application to identify objects in the images collected through web scraping. Once we identify the objects in an image we use collaborative filtering to establish a match with keywords that were previously extracted from a film description. From the analysis and performance metrics, YOLO, Mask R-CNN, and SSD are more efficient in terms of time and space and thus more suited to our problem.

5 References

- Campos, R., Mangaravite, V., Pasquali, A., Jorge, A.M., Nunes, C., Jatowt, A. (2018). YAKE! Collection-Independent Automatic Keyword Extractor. In: Pasi, G., Piwowarski, B., Azzopardi, L., Hanbury, A. (eds) Advances in Information Retrieval. ECIR 2018. Lecture Notes in Computer Science(), vol 10772. Springer, Cham. https://doi.org/10.1007/978-3-319-76941-780

- Bert-Based Text Keyword Extraction, Journal of Physics: Conference Series, Qian Yili, Jia Chaochao, Liu Yimei,2021, doi:10.1088/1742-6596/1992/4/042077.

- M. G. Thushara, T. Mownika and R. Mangamuru, "A Comparative Study on different Keyword Extraction Algorithms," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), 2019, pp. 969-973, doi: 10.1109/ICCMC.2019.8819630.

- Zhao, Zhong-Qiu Zheng, Peng Xu, Shou-Tao Wu, Xindong. (2019). Object Detection With Deep Learning: A Review. IEEE Transactions on Neural Networks and Learning Systems. PP. 1-21. 10.1109/TNNLS.2018.2876865.