Multi-document Summarization via Deep Learning Techniques: A Survey

CONGBO MA,WEI EMMA ZHANG,MINGYU GUO,HU WANG,QUAN Z. SHENG

Presentation Structure

- Introduction
- Key Contributions
- Shift from Single to Multi-Document Summarization
- Deep learning Based multi-document summarization Methods
- Objective Functions
- Evaluation Metrics
- Data sets
- Future directions

Introduction

Key Contributions

Key Contributions

- The Paper propose a categorization scheme to organize current research and provide a comprehensive review for deep learning based MDS techniques, including deep learning based models,objective functions, benchmark datasets and evaluation metrics.
- Researchers have reviewed development movements and provide a systematic overview and summary of the state-of-the-art. They also summarize nine network design strategies based on their extensive studies of the current models.
- Researchers have discussed the open issues of deep learning based multi-document summarization and identify the future research directions of this field. They also proposed potential solutions for some discussed research directions.

Shift from Single to Multi-Document Summarization

Shift from Single to Multi-Document Summarization

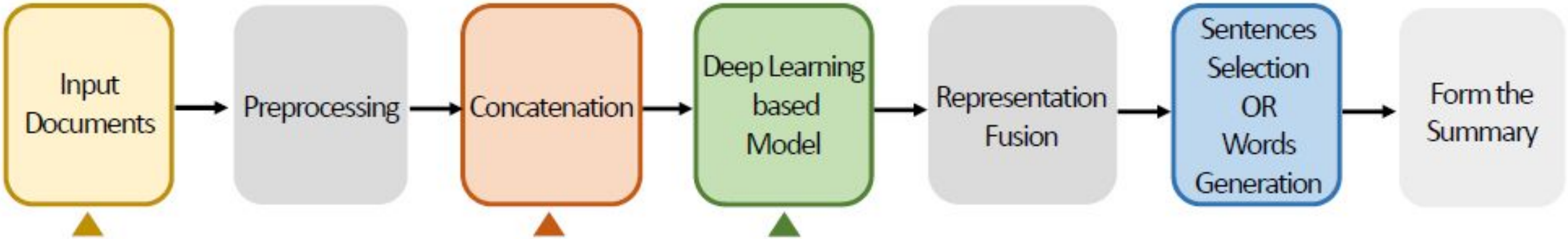

Fig1 : The Processing Framework of Text Summarization

Shift from Single to Multi-Document Summarization

Research questions of MDS via Deep Learning Techniques:

- How to capture the cross-document relations and in-document relations from the input documents?
- Compared to SDS, how to extract or generate salient information in a larger search space containing conflict, duplication and complementary information?
- How to best fuse various representation from deep learning based models and external knowledge?
- How to comprehensively evaluate the performance of MDS models?

Shift from Single to Multi-Document Summarization

Similarities between SDS and MDS:

- SDS and MDS both seek to compress the document(s) in to a short and informative summary.
- Both utilize the same summarization techniques (abstractive summarization, extractive summarization and hybrid summarization).
- both SDS and MDS aim to minimize the distance between machine-generated summary and golden summary

Shift from Single to Multi-Document Summarization

Differences between SDS and MDS:

- A defining different character between SDS and MDS is the number of input documents.
- MDS has to deal with larger searching space which contains high redundancy and contradiction across input documents.

Deep learning Based multi-document summarization Methods

Deep learning Based multi-document summarization Methods

Architecture Design Strategies:

-
- Naive Networks (Figure 2(a)).
- Ensemble Networks (Figure 2 (b)).
- Auxiliary Task Networks (Figure 2 (c)).
- Reconstruction Networks (Figure 2 (d)).
- Fusion Networks (Figure 2 (e)).
- Graph Neural Networks (Figure 2 (f)).
- Encoder-Decoder Structure (Figure 2 (g)).
- Pre-trained Language Models (Figure 2 (h))
- Hierarchical Networks (Figure 2 (i)).

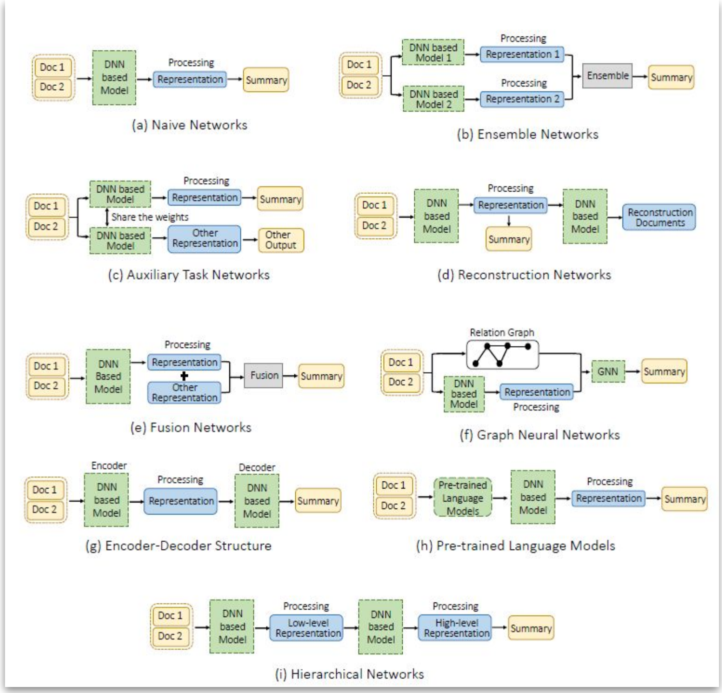

Fig 2

Deep learning Based multi-document summarization Methods

Deep Learning based Methods :

Methods	Works	Construction Types			Document-level Relationship		Comparison of DL based techniques
		Ext	Abs	Hyb	FC	HC	Pros and Cons
RNN	MeanSum [30]		✓		✓		Pros: Can capture sequential relations and syntactic/semantic information from word sequences Cons: Not easy to parallel computing; Highly depending on results from the previous steps
	Zhang et al. [171]		✓		✓		
	STDS [181]	✓				✓	
	ParaFuse_doc [111]		✓			✓	
	R2N2 [19]	✓			✓		
	CondaSum [3]			✓		✓	
	C-Attention [89]		✓		✓		
	Wang et al.[157]		✓		✓		
	RL-MMR [102]	✓			✓		
	Coavoux et al.[32]		✓		✓		
CNN	MV-CNN [178]	✓			✓		Pros: Good parallel computing; Cons: Not good at processing sequential data
	TCSum [18]	✓			✓		
	CNNLM [168]	✓			✓		
	PriorSum [20]	✓			✓		
	Angelidis et al.[4]	✓			✓		
GNN	Yasunaga et al.[167]	✓				✓	Pros: Can capture cross-document and in-document relations Cons: Inefficient when dealing with large graphs
	SemSentSum [5]	✓				✓	
	Scisummnet [166]	✓				✓	
	HDSG [155]	✓				✓	
PG	PG-MMR [83]		✓		✓		Pros: Low redundancy Cons: Hard to train
	Hi-MAP [43]		✓		✓		
Transformer	HT [94]		✓			✓	Pros: Good performance; Good parallel computing; Can capture cross-document and in-document relations Cons: Time-consuming; Problems with position encoding
	MGSum [71]	✓	✓			✓	
	FewSum [16]	✓	✓		✓		
	GraphSum [90]		✓			✓	
	Bart-Long [122]		✓			✓	
	WikiSum [93]			✓	✓		
Deep Hybid Model	Cho et al.[28]	✓			✓		Pros: Combines the advantages of different DL models Cons: Computationally intensive
	GT-SingPairMix [82]		✓		✓		
	HNet [141]	✓			✓		

Deep learning Based multi-document summarization Methods

The Variants of Multi-document Summarization :

- Query-oriented MDS
- Dialogue summarization
- Stream summarization

Objective Functions

Objective Functions

- Cross-Entropy Objective
- Reconstructive Objective
- Redundancy Objective
- Max Margin Objective
- Multi-Task Objective

Evaluation Metrics

Evaluation Metrics

Evaluation Metrics		Advantages	Disadvantages
Lexical Matching Metrics	ROUGE	• Widely used • Intuitive • Easily computed	• Cannot measure texts semantically • Exact matching
	BLEU	• Intuitive • Easily computed • High correlations with human judgments	• Cannot measure texts semantically • Cannot deal with languages lacking word boundaries
	Perplexity	• Easily computed • Intuitive	• Sensitive to certain symbols and words
	Pyramid	• High correlations with human judgments	• Requires manually extraction of units • Bias results easily
	Responsiveness	• Consider both content and linguistic quality • Can be calculated without reference	• Not widely adopted
	Data Statistics	• Can measure the density and coverage of summary	• Cannot measure texts semantically
Semantic Matching Metrics	MEREOR	• Consider non-exact matching	• Sensitive to length
	SUPERT	• Can measuring texts semantic similarity	• Not widely adopted
	Preferences based Metric	• Does not depend on the golden summaries	• Require human annotations
	BERTScore	• Semantically measure texts to some extent • Mimic human evaluation	• High computational demands
	MoverScore	• Semantically measure texts to some extent • More similar to human evaluation by adopting earth mover's distance	• High computational demands
	Importance	• Combining redundancy, relevance and informativeness • Theoretically supported	• Non-trivial for implementation
	Human Evaluation	• Can accurately and semantically measure texts	• Require human annotations

Data sets

Data sets

Datasets	Cluster #	Document #	Summ #	Ave Summ Len	Topic
DUC01	30	309 docs	60 summ	100 words	News
DUC02	59	567 docs	116 summ	100 words	News
DUC03	30	298 docs	120 summ	100 words	News
DUC04	50	10 docs / cluster	200 summ	665 bytes	News
DUC05	50	25-50 docs / cluster	140 summ	250 words	News
DUC06	50	25 docs / cluster	4 summ / cluster	250 words	News
DUC07	45	25 docs / cluster	4 summ / cluster	250 words	News
TAC 2008	48	10 docs / cluster	4 summ / cluster	100 words	News
TAC 2009	44	10 docs / cluster	4 summ / cluster	100 words	News
TAC 2010	46	10 docs / cluster	4 summ / cluster	100 words	News
TAC 2011	44	10 docs / cluster	4 summ / cluster	100 words	News
OPOSUM	60	600 rev	1 summ / cluster	100 words	Amazon reviews
WikiSum	-	train / val / test 1579360 / 38144 / 38205	1 summ / cluster	139.4 tokens/ summ	Wikipedia
Multi-News	-	train / val / test 44972 / 5622 / 5622 2-10 docs / cluster	1 summ / cluster	263.66 words / summ 9.97 sents / summ 262 tokens / summ	News
Opinosis	51	6457 rev	5 summ / cluster	-	Site reviews
Rotten Tomatoes	3731	99.8 rev / cluster	1 summ / cluster	19.6 tokens / summ	Movie reviews
Yelp	-	train / val / test bus: 10695 / 1337 / 1337 rev: 1038184 / 129856 / 129840	-	-	Customer reviews
Scisumm	1000	21 - 928 cites / paper 15 sents / refer	1 summ / cluster	151 words	Science Paper
WCEP	10200	235 docs / cluster	1 summ / cluster	32 words	Wikipedia
Multi-XScience	-	train / val / test 30369 / 5066 / 5093	1 summ / cluster	116.44 words / summ	Science Paper

Future directions

Future directions

- Creating More High-quality Datasets for MDS
- Improving Evaluation Metrics for MDS
- Pre-trained Language Models for MDS
- Creating Explainable Deep Learning Model for MDS
- Adversarial Attack and Defense for MDS

Q & A