A REVIEW ON NEURAL NETWORK QUESTION ANSWERING SYSTEMS

Lorena Kodra and Elinda Kajo Meçe

Department of Computer Engineering, Polytechnic University of Tirana, Albania

ABSTRACT

In recent years neural networks (NN) are being used increasingly on Question Answering (QA) systems and they seem to be successful in addressing different issues and challenges that these systems exhibit. This paper presents a review to summarize the state of the art in question answering systems implemented using neural networks. It identifies the main research topics and considers the most relevant research challenges. Furthermore, it analyzes contributions, limitations, evaluation techniques, and directions proposed for future research.

KEYWORDS

Information Retrieval, Question Answering Systems, Neural Networks

1. INTRODUCTION

In recent years neural networks are being used increasingly on QA systems and have been proven effective in achieving good results [45]. In this paper we present a study of the latest research being done on question answering systems based on neural networks. Our aim is to give an answer to questions like: What are the research directions of Neural Network Question Answering? What are the biggest challenges that researchers face in this area? What are the solutions proposed? What criteria and datasets are used for evaluations? Which are the most promising directions for future research? Our work can help researchers gain an insight on the state of the art and future research directions of the research being done in the area of neural network question answering systems. The rest of this paper is organized as follows: In section 2 we describe the methodology used in our study and define objectives and research questions. Section 3 outlines the results of the review organized according to each research question defined in Section 2. Section 4 discusses the results of our study. Section 5 contains conclusions and future work.

2. METHODOLOGY

2.1. RESEARCH QUESTIONS

As a primary step in the investigation, retrieval and selection of the most accurate publications for our review we have defined the following research questions:
RQ1: What are the research directions of neural network question answering?
RQ2: What are the challenges faced by researchers in this area?
RQ3: What solutions have been proposed?
RQ4: What criteria and datasets are used for evaluations?
RQ5: Which directions are the most promising for future research?

2.2. SEARCH KEYWORDS AND SOURCE SELECTION

In order to extract the most relevant information for our review we used the following keywords and their combination and synonyms. The search string below was used as a query to search for publications in different online digital libraries: ("question answering" OR "question answer" OR "question answering system" OR "question answering systems").

The search string contained only these keywords because not all the publications had on their title the keywords "neural" or "neural network". After retrieving the full publication another search was made on the body of the publication with the keyword "neural". If the search was successful and neural networks were identified as an implementation mechanism the publication was included in the list of publications.

Furthermore, we selected four scientific digital libraries that represent primary sources for computer science research publications. We did not include online archives Google Scholar and ArXiv because they index content from existing digital libraries. The sources are shown in Table 1.

Table 1. Selected sources for the search process.

Source	URL
IEEExplore	http://ieeexplore.ieee.org
ScienceDirect	http://www.sciencedirect.com/
ACM Digital Library	http://dl.acm.org
Springer Link	http://link.springer.com

2.3. INCLUSION AND EXCLUSION CRITERIA

Table 2 lists the inclusion and exclusion criteria that we used to collect publications.

Table 2. Inclusion and exclusion criteria.

Inclusion criteria	Exclusion criteria
Relevant to the topic of our review	Review papers
Papers that have been published in the last three years (2014 - 2016)	Reports
Papers from conferences and journals	Papers reporting only abstract or presentation slides because of the lack of information describing the corresponding study.
Short and workshop papers that fulfil the study criteria	Grey literature. PhD thesis, unpublished studies and technical reports would not add much more information against journal and conference papers.
English language	

We did not retrieve review papers and reports because our aim is to analyze the existing implementations and latest developments of Question Answering Systems (QAS) based on neural networks.

3. RESULTS

This section presents the relevant information found in the selected publications in order to answer the proposed research questions. A further discussion and analysis of these results are addressed in Section 4.

3.1. RESEARCH DIRECTIONS

We retrieved a total of 46 publications to include in our review. We identified three research directions as shown in Table 3.

Table 1. Distribution of publications according to research directions

Research direction	Number of publications
Knowledge Based Question Answering	18
Visual Question Answering	15
Community Question Answering	13

Knowledge Base Question Answering (KBQA) is an area that deals with answering questions over a Knowledge Base. In such systems the information is organized in a form of structured knowledge base (KB) where the data are linked by semantics. The information in the knowledge base is organized in the form of triples composed of subject, predicate, and object. An example of a triple is (Mount Everest, height, 8.848 m). The predicate is a characteristic of the subject, while the object is the value of that characteristic. In question answering over a knowledge base, the question is on subject and predicate, while the answer is from object.

Visual Question Answering (VQA) is an area that deals with answering questions about an image. The information source is either a hybrid between text and image or entirely image-based.

Community Question Answering (CQA) Systems are modern QA systems where users rely on expertize from the community to get an answer for their question. Examples of CQAs include Yahoo! Answers[1] and Quora[2].

For each of the research directions we identified the main research topics as shown in tables 4, 5 and 6.

Table 2. Distribution of publications according to research topics in Knowledge Based Question Answering

Research topic	Number of publications
Entity linking	6
Answer selection	5
Question-answer matching	4
Question answering over paragraph questions	2
Learning to translate in multilingual QAS	1
Natural language answering	1

---

[1] https://answers.yahoo.com/

[2] https://www.quora.com/

Table 3. Distribution of publications according to research topics in Visual Question Answering

Research topic	Number of publications
Answering free language questions about an image	7
Identifying clues in image and/or question to generate answer	4
Detecting relevance of question to image.	1
Visual analogy	1
Situated visual question answering	1
Answering from a limited domain of questions	1

Table 4. Distribution of publications according to research topics in Community Question Answering

Research topic	Number of publications
Retrieving similar questions	5
Question-answer matching	3
Expert finding	2
Answer ranking	2
Answer selection	2

3.2. RESEARCH CHALLENGES

In order to address RQ2 we summarize the main research challenges involved in the selected publications. Tables 7 and 8 present these challenges for KBQAS and CQAS respectively. As regards VQAs, we consider challenges to be the same with research topics.

Table 5. Distribution of publications according to research challenges in Knowledge Based Question Answering

Research challenges	Number of publications
Lexical gap between natural language and structured semantics of the knowledge base	10
Entity identification and linking	5
Multiple entity question	1
Natural language answer generation	1
Questions involving multiple entities	1
Passage question answering	1

Table 6. Distribution of publications according to research challenges in Community Question Answering

Research challenges	Number of publications
Lexical gap between questions	5
Lexical gap between questions and answers	5
Sparsity of CQA data	2
Deviation from question	2

3.2.1. RESEARCH CHALLENGES IN KNOWLEDGE BASE QUESTION ANSWERING SYSTEMS

We identified the following challenges for KBQA systems:

- **Lexical gap between natural language and structured semantics of the knowledge base:** We identified it as the most frequent problem. It concerns differences in sentence

representations between the unstructured natural language question and the structured knowledge base. It also concerns the many ways of expressing knowledge in a knowledge base.

- **Entity identification and linking:** This was another prominent challenge. It concerns the ability of the system to correctly identify the subject entity in question and link it to a triple in the knowledge base.

- **Natural language answer generation:** It concerns the ability of a system to generate the answer in a natural language form and not simply give the information user asked for. For example for the question "How many days are there in a leap year?" the system may answer "There are 366 days in a leap year." instead of just "366". This feature is desirable in making the system seem "smarter".

- **Questions involving multiple entities:** It concerns the ability of the system to identify and reason over multiple subject entities in question and link it to the relevant triple in the knowledge base.

- **Passage question answering:** This is a challenge on non-factoid question answering where the answer is in the form of a paragraph. Question-answer matching is a challenging task as it requires effective representations that capture the complex semantic relations between questions and answers.

3.2.2. RESEARCH CHALLENGES IN VISUAL QUESTION ANSWERING

We identified the following challenges for VQA systems:

- **Answering free language questions about an image:** We identified it as the most frequent challenge. It concerns the ability of the system to answer questions posed in a free language (i.e. without any lexical or semantic limitation). It is a complex research problem since the system is required to answer different types of questions ranging from simple questions like "What is the color of the flower?" to "Does it appear to be rainy?" which requires extra knowledge or commonsense reasoning.

- **Identifying clues in image and/or question to generate answer:** It was another frequent challenge. It concerns the ability of the system to reason about question and/or image to identify possible words or image regions that can be helpful in answering the question.

- **Visual analogy:** This topic deals with finding analogue images from a given set. The questions take the form of "image A is to image B as image C is to what". Answering these questions requires discovering the mapping from image A to image B, extending this mapping to image C and finding image D so that the relation from A to B holds for C to D.

- **Situated visual question answering:** It concerns the problem of answering questions about an environment such as an image or diagram. This problem requires jointly reasoning about uncertain interpretations of both a question and an environment together with background knowledge to select the correct answer.

- **Answering from a limited domain of questions:** This topic concerns systems that can accept simple factoid question. The answer also may be in a form of a single word. Sometimes the questions can be only from a single domain (e.g. health).

3.2.3. RESEARCH CHALLENGES IN COMMUNITY QUESTION ANSWERING

We identified the following challenges for CQA systems:

- **Lexical gap between questions:** It was one of the most frequent problems in the selected publications. It concerns differences in natural language formulation of questions. Different users ask for the same information but they formulate the question in different ways. This results in many questions that are semantically equivalent but differ lexically.

- **Lexical gap between questions and answers:** This was another frequent problem. Similar to the lexical gap between questions, sometimes question and answers can be highly asymmetric in the information they contain. There is also a technical terminology gap between questions and answers. Questions are posed by novices or non-experts who use less technical terminology while experts who answer questions use the correct terms.

- **Sparsity of CQA data:** Another problem that we identified is the low participation of answerers. Usually, each question is answered by a few users. As a consequence the system cannot identify relevant answerers and most new questions cannot be routed to appropriate users who can answer them.

- **Deviation from question:** It concerns the phenomenon of answer thread becoming irrelevant to the question. Answers are given in the form of comments but sometimes users engage in discussion and deviate from the original question.

3.3. PROPOSED SOLUTIONS

In order to address RQ3 we present and classify the solutions provided to research problems for each research direction. In some cases there is only one publication dealing with a solution. Due to the differences among solutions we cannot group or merge them into a single category.

3.3.1. PROPOSED SOLUTIONS IN KNOWLEDGE BASE QUESTION ANSWERING SYSTEMS

Lexical gap between natural language and structured semantics of the knowledge base:
Different solutions have been proposed to close this lexical gap. Some of them share the common concept of using neural networks to obtain the semantic similarity by learning a similarity metric. For example, for the task of answer selection a neural network is used to analyze question-answer pairs from different aspects (i.e. answer path, answer context, and answer type) and then computes a similarity score for the pair. The system uses probabilistic inference to select candidates with the best scores [15]. Another proposed solution is a neural network that sequentially reads the words from question and answer sentences and then outputs their relevance scores. In order to identify the correct candidate answer sentences the system does keyword matching between the words of the question and answer [18] [31].

For the task of question-answer matching the proposed solutions that we identified include a neural network that uses graph-based semantic parsing to extract meaning, entities and relations by leveraging the knowledge base more tightly when parsing input questions by linking entities in the question to entities in the knowledge base [16], Or combination of semantic and structural features to compute sentence similarity [42]. Another approach proposed is to look for answers

that are related to the question in a relevant way, according to the information need of the question which may be determined through task-specific word embedding (i.e. different embedding for different types of questions) [27]. A study using a factual memory network, which learns to answer questions by extracting and reasoning over relevant facts from a knowledge base, is also presented. Facts are visualized as a question-answer pair and are generated by matching all n-grams of words of the question against knowledge base entities. To find the relevant answer, the neural network calculates the similarity between the actual question and a hypothetical question constructed from the fact [25].

For the task of entity linking, neural networks are used to measure the similarity of entity mentions in the question with entities in the knowledge base. Triples in the knowledge base are scored using these similarity measures and the top scoring triple is select to answer the question [20], [30].

For multilingual systems, a proposed solution for the task of translating question into knowledge base language or vice-versa is modeled by a neural network that learns the optimal translation of question and/or candidate answer, based on how well it discriminates between good and bad answers [28].

ENTITY IDENTIFICATION AND LINKING:

There were various solutions proposed to tackle this problem. One of them was to use a multistep approach to "zooming in" a question to find more probable candidate subject mentions by using a neural network to first identify focus segments in a question that are more probable of containing the subject entity and then pruning the candidate knowledge base triples according to textual similarities between question segment and knowledge base triple subject [9]. Another solution was to use a neural network to compute character-level similarity score between the fact candidate and entity in the question to determine the correct entity. Then use a word-level match (pattern match) similarity score between the predicate in the fact and question words to find the correct predicate [21]. Another proposed alternative was to use a neural network to tag whether each question word is part of an entity, and optionally further divide into subject, object, and predicate classes. Then converting the input question into a structured query, which can be ran against the knowledge graph to retrieve the fact(s) that answer(s) the question [37]. In the case of paragraph questions, the solutions proposed were to use a neural network to learn word and phrase-level representations that combine across question sentences to capture the topic of the paragraph and reason about entities [33], [43].

NATURAL LANGUAGE ANSWER GENERATION:

We identified only one publication trying to solve this problem and the proposed solution was a generative neural network system that takes a sequence of words as input question and generates another sequence of words as output answer. To provide right answers, the system is connected with a knowledge base that contains facts. During the process of answering, the system queries the knowledge base, retrieves a set of candidate facts and generates a correct answer to the question using the right fact. The generated answer may contain two types of "words": one is common words for composing the answer (referred to as common word) and the other is specialized words in the knowledge base denoting the answer (referred to as knowledge base word). The generated answer is in the form of a whole sentence in natural language instead of a simple fact [1].

QUESTIONS INVOLVING MULTIPLE ENTITIES:

The only publication that reported this research problem proposed a relation extractor neural network that selects answer candidates from the knowledge base and then infers over external unstructured sources of evidence like Wikipedia[3] to choose the correct answer [11].

PASSAGE QUESTION ANSWERING:

The only publication that dealt with this research challenge proposed a non-factoid question answering system that used neural networks to extract linguistic information from both question and passage answer. Attention mechanisms where proposed to generate answer representations according to the question, such that the embeddings do not overlook informative parts of the answers [10].

Table 9 summarizes the proposed approaches mentioned above. Some systems use more than one method.

Table 7. Distribution of publications according to methods used for Knowledge Base Question Answering

Method	Number of publications
Similarity measures	8
Representation learning	6
Ranking mechanisms	3
Inference	3

3.3.2. PROPOSED SOLUTIONS IN VISUAL QUESTION ANSWERING SYSTEMS

ANSWERING FREE LANGUAGE QUESTIONS ABOUT AN IMAGE:

Various alternatives have been proposed regarding this research topic. Some of them use a common idea such as attention mechanisms. The authors in [24] propose an attention model for answering questions about a variety of world representations, including images and structured knowledge bases. The model uses attention mechanisms to decompose the answering process into modular sub-problems. It translates from questions to dynamically assembled neural networks based on the type of question, then applies these networks to world representations (images or knowledge bases) to produce answers. Another attentional system is proposed in [35]. This system is composed of a stacked attention net (SAN) that uses semantic representation of a question as a query to identify the regions of the image that are related to the answer. The SAN is multi-layered and it queries the image multiple times to infer the answer progressively. The authors state that the motivation behind this approach is the fact that image question answering often requires multiple steps of reasoning. For example for the question: "What is the girl holding? The system must first identify the girl. Then the system must identify her hands. Then it needs to understand what the girl is holding.

The authors in [5] propose a reasoning model that is based on neural networks and is able to update the question representation iteratively by inferring image information. The model updates the question representation iteratively by selecting image regions relevant to the query and learns to give the correct answer. A neural reasoned updates the question by interacting it with supporting facts through multiple reasoning layers. With this technique, it is possible to make questions more specific than the original ones focusing on important image information automatically.

---

[3] https://www.wikipedia.org/

Another proposed solution is a framework of neural networks for learning not only the image and question representations, but also their inter-modal interactions to produce the answer [39].
Another alternative are multimodal systems composed of different neural network types to extract question information, visual representation, store the linguistic con-text of the answer and combine this information into generating a relevant answer to a free language question have also been proposed by [7], [26] and [36].

IDENTIFYING CLUES IN IMAGE AND/OR QUESTION TO GENERATE ANSWER:

The authors in [4] use attention mechanisms to answer questions by finding both "where to look" in an image as well as "what words to listen to" in a question. The mechanism proposed, jointly reasons about visual attention, by identifying regions in an image that are relevant to the question, and question attention, by identifying words in a question that might contain useful information for answering the question, referred to as co-attention, which has a natural symmetry between the image and question. It reasons about the questions (and consequently about the image via the co-attention mechanism) in a hierarchical fashion by word level, phrase level and question level. Other attention-based systems proposed in [44] and [47] use neural networks to store into memory information about different regions of the image relevant for computing the answer.
Another proposed solution we identified is a dynamic parameter neural network whose parameters are determined adaptively based on input questions [34]. In this way the system reasons differently for each question. The authors stated that the motivation behind this approach is the fact that different questions require different types and levels of understanding of an image to find correct answers. For example, to answer the question "is this an apple?" we should decide between yes and no, while for questions like "how is the weather?" we need to perform classification on multiple choices related to weather. Hence, besides performance on a single recognition task, the capability to select a proper task is equally important.

VISUAL ANALOGY:

We identified one publication dealing with this topic. In this study, the analogy problem is treated as learning an embedding that makes pairs of analogous images with similar transformations to be close together in the embedding space and pushing apart images with dissimilar transformations [6].

SITUATED VISUAL QUESTION ANSWERING:

We identified one publication dealing with situated visual question answering. The authors formulated situated question answering as semantic parsing with an execution model that is a learned function of the environment that can use background knowledge and global features of the question/environment interpretation while retaining efficient approximate inference to answer a question [29].

ANSWERING FROM A LIMITED DOMAIN OF QUESTIONS:

The only publication that we identified for this research topic assumes that the answer consists of only a single word and treats the answering process as a classification process. The system treats the image as a word of a sentence. The authors propose the use of a neural network and visual semantic embeddings without intermediate stages like object detection and image segmentation to predict answers to simple questions about images [8].

DETECTING RELEVANCE OF QUESTION TO IMAGE:

We identified one study dealing with this research challenge. The proposed system uses a neural network to determine the relevance of questions to images in a 2-step fashion. First it captures visual-specific linguistic structure to determine if the question is visual or not. If the question is visual, pre-trained captioning models generate relevant captions (or questions) for the given image and then compare them to the given question to determine if the visual question is relevant to the image [32]

Table 10 summarizes the proposed approaches mentioned above and illustrates the distribution of publications in accordance with the proposed approaches.

Table 8. Distribution of publications according to approaches proposed for Visual Question Answering

Approach	Number of publications
Attention mechanisms	6
Multimodal systems	3
Dynamic parameter NN	1
Other	5

3.3.3. PROPOSED SOLUTIONS IN VISUAL QUESTION ANSWERING SYSTEMS

LEXICAL GAP BETWEEN QUESTIONS:

Various solutions have been proposed to overcome this problem. They share the common concept of using neural networks to obtain the semantic similarity by learning a similarity metric. The solutions include capturing the semantic similarity be-tween the current and archived questions and incorporating similarity scores to rank and select semantically equivalent questions [12], [17], [22], also incorporating metadata information such as question category to learn the similarity metric [14]. Another proposed solution was to use a neural network to first learn semantic representation of question and answer pairs from a collection of question and answer pairs. Then for each pair use a mixture score to compute its similarity with queried question [3].

LEXICAL GAP BETWEEN QUESTIONS AND ANSWERS:

This problem is relevant to the task of matching question and answer. The proposed solutions we identified are similar to the problem of lexical gap between questions. They include using neural networks to learn the similarity metric between questions and answers and incorporating similarity scores to rank and select the best match [40], [48], and also enhance the quality of question-answer matching by collaboratively utilizing the rich interaction among questions, answers and answerers to learn the relative quality rank of different answers with a same question [38].

SPARSITY OF CQA DATA:

The problem of data sparsity is relevant to the task of expert finding. A proposed solution to tackle this problem was modeling the expertise of a user based on his/her answering history by combining the questions answered by the user for which he or she has been selected as the best answerer. A neural network is used to predict best answerers for new questions by classifying the new questions and then obtaining the probabilities of every user to be the best answer [46]. Another proposed solution was using a neural network to learn the ranking metric by leveraging both user answering history and their social relation with other experts in the community with the

idea that a social relation between two users provides a strong evidence for them to have common background knowledge [2].

DEVIATION FROM QUESTION:

This problem is relevant to the task of answer selection and ranking. The proposed solutions identified for this problem were borrowing ideas from Machine Translation Evaluation (MTE) [41] by using a neural network to decide the quality of an answer comment by taking two comments and use ranking mechanisms to filter out bad (irrelevant) comments [13], and using a neural network to learn the joint semantic representation of a question-answer pair and use this representation to predict the quality of each answer in the comments sequence [19].

Table 11 summarizes the proposed methods mentioned above. Some systems use more than one method.

Table 9. Distribution of publications according to methods used to solve Community Question Answering

Methods	Number of publications
Similarity metric learning	8
Probabilistic inference	6
Ranking mechanisms	6

3.4. EVALUATION TECHNIQUES

In order to address RQ4 we identified and collected data about the datasets used for evaluating the proposed studies, as well as metrics used in the evaluation. We find worth mentioning that some of the publications used multiple datasets and multiple metrics for evaluation. Table 12 presents the total distribution of publications according to the datasets they use for evaluation.

Table 10. Distribution of publications according to the dataset used for evaluation

Dataset	Number of publications
Self-constructed	14
VQA	9
COCO-QA	8
DAQUAR	7
Yahoo! Answers	6
WebQuestions	4
Baidu Zhidao	3
SIMPLE-QUESTIONS	3
AskUbuntu	2
InsuranceQA	1
TREC-QA	1
SemEval 2015 CQA	1
PARALAX	1
GeoQA	1
Visual7W	1
InfoboxQA	1
BabI-10k English	1
StackOverflow	1

We notice that the most used datasets are self-constructed, accounting for 30% of total publications. Among the publicly available datasets, the most used ones are visual datasets VQA, COCO-QA, DAQUAR and Yahoo! Answers, WebQuestions, Baidu Zhidao and Simple-Questions.

Table 13 presents the total distribution of publications according to the metrics they use for evaluation.

Table 11. Distribution of studies according to evaluation metrics

Metric	Number of publications
Accuracy	23
MRR	9
F1	7
MAP	6
P@K	6
WUPS	5
NDCG, DCG@K	4
Recall	3
Precision	2
Fluency	1
AvgRec	1
R-Prec	1
SVMRank	1
S@N	1
Visual Turing Test	1

3.5. FUTURE RESEARCH DIRECTIONS

To address RQ5 we analyzed the future work that the selected publications proposed in order to improve or extend their approaches. We also analyzed the current system limitations. The information collected is summarized in tables 14, 15 and 16.

Table 12. Future research directions for KBQA

Future research directions
Extend system to handle more complex questions
Include support for compositional questions or multiple entity questions
Integrate external knowledge sources
Extend model for other tasks like CQA
Add support for KB with noisy data or unstructured KB
Include non-factoid questions
Include larger dataset for testing
Improve accuracy

Table 13. Future research directions for CQA

Future research directions
Enhance system to include meta-data information like user votes, rating, user reputation, etc
Incorporate more similarity measures
Improve performance
Improve topic recognition
Apply more sophisticated NN models

Table 14. Future research directions for Visual QA

Future research directions
Improve object-counting ability
Improve word-region relation
Add support for longer answers
Include visual attention
Expand domain of question
Add support for non-visual questions by leveraging an external knowledge
Improve accuracy

4. DISCUSSION

The purpose of this paper was to make a quantitative and topic-based analysis of the recent developments of QAS implemented with neural networks. The scope of this study was not to give a detailed comparison between different NN approaches in terms of performance or compare them with traditional approaches.

To answer RQ1 we identified current research directions of Neural Network Question Answering. The research directions are Knowledge Base Question Answering, Visual Question Answering and Community Question Answering. Knowledge Base Question Answering has the most number of contributions with 39.1% of the total. However there is no considerable difference in the quantity of research contributions among the three directions. VQA has 32.7% of total contributions and CQA 28.2%. Therefore we cannot say that KBQA is the prevalent research direction.

Regarding RQ2 we can say that the main challenges faced today by researches are: 1) lexical gap between natural language questions and answers, as well as structured semantics of the knowledge base for text based question answering, 2) answering free language questions about an image due to the fact that the system must reason differently for different kind of questions for image based question answering. These challenges clearly have the most number of contributions. To answer RQ3 we can say that for text based question answering the most used methods and techniques are similarity measures, representation learning and inference while for image based question answering the most used approaches are attention mechanisms and multimodal systems. RQ4 regards datasets and metrics used to evaluate the proposed system. From the results presented in Section 3 we noticed that a considerable number of studies used self-constructed datasets to evaluate the proposed solutions. Among the publicly available datasets, the most used ones are VQA, COCO-QA, DAQUAR and Yahoo! Answers, WebQuestions, Baidu Zhidao and Simple-Questions. Regarding evaluation metrics we can say that the most used metrics are Accuracy, Mean Reciprocal Rank (MRR) and F1.

Regarding RQ5, common future directions proposed for all types of systems include: improve accuracy, increase the complexity of supported questions and answers and integrate external knowledge sources. As QAS are considered to be the future of search engines, these improvements are justified by the need to make QAS more interactive, more intelligent and closer to humans.

5. CONCLUSIONS

In this paper we presented a study on the state of the art of the latest research being done on QAS based on neural networks. We have identified the main research directions and answered the defined research questions regarding them. We have presented the main challenges faced by researchers in this area along with the proposed solutions and future research directions. Evaluation criteria and datasets used for testing were also identified. From this study we can conclude that NN are being used effectively and continuously improving their performance, but still there are unsolved problems and new challenges arising. As a future work we plan to test publicly available QAS on common datasets and metrics relevant to their type and evaluate them.

REFERENCES

[1] Yin, Jun & Jiang, Xin & Lu, Zhengdong & Shang, Lifeng & Li, Hang & Li, Xiaoming (2016) "Neural Generative Question Answering", IJCAI (International Joint Conference on AI)

[2] Zhao, Zhou & Yang, Qifan & Cai, Deng & He, Xiaofei & Zhuang, Yueting (2016) "Expert Finding for Community-Based Question Answering via Ranking Metric Network Learning", IJCAI (International Joint Conference on AI)

[3] Qiu, Xipeng & Huang, Xuanjin (2015) "Convolutional Neural Tensor Network Architecture for Community-Based Question Answering", IJCAI (International Joint Conference on AI)

[4] Lu, Jiasen & Yang, Jianwei & Batra, Dhruv & Parikh, Devi (2016) "Hierarchical Question-Image Co-Attention for Visual Question Answering", NIPS (Neural Information Processing Systems Conference)

[5] Li, Ruiyu & Jia, Jiaya (2016) "Visual Question Answering with Question Representation Update", NIPS (Neural Information Processing Systems Conference)

[6] Sadeghi, Fereshteh & Zitnick, C. Lawrence & Farhadi, Ali (2015) "Visalogy: Answering Visual Analogy Questions", NIPS (Neural Information Processing Systems Conference)

[7] Gao, Haoyuan & Mao, Junhua & Zhou, Jie & Huang, Zhiheng & Wang, Lei & Wei, Xu (2015) "Are You Talking to a Machine? Dataset and Methods for Multilingual Image Question", NIPS (Neural Information Processing Systems Conference)

[8] Ren, Mengye & Kiros, Ryan & Zemel, Richard S. (2015) "Exploring Models and Data for Image Question Answering", NIPS (Neural Information Processing Systems Conference)

[9] Dai, Zihang & Li, Lei & Xu, Wei (2016) "CFO: Conditional Focused Neural Question Answering with Large-scale Knowledge Bases", ACL (Annual Meeting of the Association on Computational Linguistics)

[10] Tan, Ming & dos Santos, Cicero & Xiang, Bing & Zhou, Bowen (2016) "Improved Representation Learning for Question Answer Matching", ACL (Annual Meeting of the Association on Computational Linguistics)

[11] Xu, Kun & Reddy, Siva & Feng, Yansong & Huang, Songfang & Zhao, Dongyan (2016) "Question Answering on Freebase via Relation Extraction and Textual Evidence", ACL (Annual Meeting of the Association on Computational Linguistics)

[12] Das, Arpita & Yenala, Harish & Chinnakotla, Manoj & Shrivastava, Manish (2016) "Together we stand: Siamese Networks for Similar Question Retrieval", ACL (Annual Meeting of the Association on Computational Linguistics)

[13] Guzmán, Francisco & Màrquez, Lluís & Nakov, Preslav (2016) "Machine Translation Evaluation Meets Community Question Answering", ACL (Annual Meeting of the Association on Computational Linguistics)

[14] Zhou, Guangyou & He, Tingting & Zhao, Jun & Hu, Po (2015) "Learning Continuous Word Embedding with Metadata for Question Retrieval in Community Question Answering", ACL (Annual Meeting of the Association on Computational Linguistics)

[15] Dong, Li & Wei, Furu & Zhou, Ming & Xu, Ke (2015) "Question Answering over Freebase with Multi-Column Convolutional Neural Networks", ACL (Annual Meeting of the Association on Computational Linguistics)

[16] Yih, Wen-tau & Chang, Ming-Wei & He, Xiaodong & Gao, Jianfeng (2015) "Semantic Parsing via Staged Query Graph Generation: Question Answering with Knowledge Base", ACL (Annual Meeting of the Association on Computational Linguistics)

[17] Dos Santos, Cícero & Barbosa, Luciano & Bogdanova, Dasha & Zadrozny, Bianca (2015) "Learning Hybrid Representations to Retrieve Semantically Equivalent Questions", ACL (Annual Meeting of the Association on Computational Linguistics)

[18] Wang, Di & Nyberg, Eric (2015) "A Long Short-Term Memory Model for Answer Sentence Selection in Question Answering", ACL (Annual Meeting of the Association on Computational Linguistics)

[19] Zhou, Xiaoqiang & Hu, Baotian & Chen, Qingcai & Tang, Buzhou & Wang, Xiaolong (2015) "Answer Sequence Learning with Neural Networks for Answer Selection in Community Question Answering", ACL (Annual Meeting of the Association on Computational Linguistics)

[20] Yih, Wen-tau & He, Xiaodong & Meek, Christopher (2014) "Semantic Parsing for Single-Relation Question Answering", ACL (Annual Meeting of the Association on Computational Linguistics)

[21] Yin, Wenpeng & Yu, Mo & Xiang, Bing & Zhou, Bowen & Schütze, Hinrich (2016) "Simple Question Answering by Attentive Convolutional Neural Network", COLING (International Conference on Computational Linguistics)

[22] Bogdanova, Dasha & dos Santos, Cícero & Barbosa, Luciano & Zadrozny, Bianca (2015) "Detecting Semantically Equivalent Questions in Online User Forums", CONLL (Conference on Natural Language Learning)

[23] Bogdanova, Dasha & Foster, Jennifer (2016) "This is how we do it: Answer Reranking for Open-domain How Questions with Paragraph Vectors and Minimal Feature Engineering", NAACL (North American chapter of ACL)

[24] Andreas, Jacob & Rohrbach, Marcus & Darrell, Trevor & Klein, Dan (2016) "Learning to Compose Neural Networks for Question Answering", NAACL (North American chapter of ACL)

[25] Jain, Sarthak (2016) "Question Answering over Knowledge Base using Factual Memory Networks", NAACL (North American chapter of ACL)

[26] Fukui, Akira & Huk Park, Dong & Yang, Daylen & Rohrbach, Anna & Darrell, Trevor & Rohrbach, Marcus (2016) "Multimodal Compact Bilinear Pooling for Visual Question Answering and Visual Grounding", EMNLP (Conference on Empirical Methods on Natural Language Processing)

[27] Sharp, Rebecca & Surdeanu, Mihai & Jansen, Peter & Clark, Peter & Hammond, Michael (2016) "Creating Causal Embeddings for Question Answering with Minimal Supervision", EMNLP (Conference on Empirical Methods on Natural Language Processing)

[28] Ture, Ferhan & Boschee, Elizabeth (2016) "Learning to Translate for Multilingual Question Answering", EMNLP (Conference on Empirical Methods on Natural Language Processing)

[29] Krishnamurthy, Jayant & Tafjord, Oyvind & Kembhavi, Aniruddha (2016) "Semantic Parsing to Probabilistic Programs for Situated Question Answering", EMNLP (Conference on Empirical Methods on Natural Language Processing)

[30] Golub, David & He, Xiaodong (2016) "Character-Level Question Answering with Attention", EMNLP (Conference on Empirical Methods on Natural Language Processing)

[31] Morales, Alvaro & Premtoon, Varot & Avery, Cordelia & Felshin, Sue & Katz, Boris (2016) "Learning to Answer Questions from Wikipedia Infoboxes", EMNLP (Conference on Empirical Methods on Natural Language Processing)

[32] Ray, Arijit & Christie, Gordon & Bansal, Mohit & Batra, Dhruv & Parikh, Devi (2016) "Question Relevance in VQA: Identifying Non-Visual And False-Premise Questions", EMNLP (Conference on Empirical Methods on Natural Language Processing)

[33] Iyyer, Mohit & Boyd-Graber, Jordan & Claudino, Leonardo & Socher, Richard & Daumé III, Hal (2014) "A Neural Network for Factoid Question Answering over Paragraphs", EMNLP (Conference on Empirical Methods on Natural Language Processing)

[34] Noh, Hyeonwoo & Seo, Paul Hongsuck & Han, Bohyung (2016) "Image Question Answering Using Convolutional Neural Network With Dynamic Parameter Prediction", Conference on Computer

Vision and Pattern Recognition, CVPR (29th IEEE Conference on Computer Vision and Pattern Recognition)

[35] Yang, Zichao & He, Xiaodong & Gao, Jianfeng & Deng, Li & Smola, Alex (2016) "Stacked Attention Networks for Image Question Answering", CVPR (29th IEEE Conference on Computer Vision and Pattern Recognition)

[36] Malinowski, Mateusz & Rohrbach, Marcus & Fritz, Mario (2015) "Ask Your Neurons: A Neural-Based Approach to Answering Questions about Images", ICCV (International Conference on Computer Vision)

[37] Ture, Ferhan & Jojic, Oliver (2016) "Ask Your TV: Real-Time Question Answering with Recurrent Neural Networks", SIGIR (Special Interest Group on Information Retrieval)

[38] Fang, Hanyin & Wu, Fei & Zhao, Zhou & Duan, Xinyu & Zhuang, Yueting (2016) "Community-Based Question Answering via Heterogeneous Social Network Learning", AAAI Conference on Artificial Intelligence

[39] Ma, Lin & Lu, Zhengdong & Li, Hang (2016) "Learning to Answer Questions from Image Using Convolutional Neural Network", AAAI Conference on Artificial Intelligence

[40] Shen, Yikang & Rong, Wenge & Sun, Zhiwei & Ouyang, Yuanxin & Xiong, Zhang (2015) "Question/Answer Matching for CQA System via Combining Lexical and Sequential Information", AAAI Conference on Artificial Intelligence

[41] Sun, Huan & Ma, Hao & He, Xiaodong & Yih, Wen-tau & Su, Yu & Yan, Xifeng (2016) "Table Cell Search for Question Answering", WWW (World Wide Web Conference)

[42] Ruan, Haipeng & Li, Yuan & Wang, Qinglin & Liu, Yu (2016) "A Research on Sentence Similarity for Question Answering System Based on Multi-feature Fusion", International Conference on Web Intelligence, WI (International Conference on Web Intelligence)

[43] Zheng, Suncong & Bao, Hongyun & Zhao, Jun & Zhang, Jie & Qi, Zhenyu & Hao, Hongwei (2015) "A Novel Hierarchical Convolutional Neural Network for Question Answering over Paragraphs", WI (International Conference on Web Intelligence)

[44] Xiong, Caiming & Merity, Stephen & Socher, Richard (2016) "Dynamic Memory Networks for Visual and Textual Question Answering", ICML (International Conference on Machine Learning)

[45] Yang, Liu & Ai, Qingyao & Spina, Damiano & Chen, Ruey-Cheng & Pang, Liang & Croft, W. Bruce & Guo, Jiafeng & Scholer, Falk (2016) "Beyond Factoid QA: Effective Methods for Non-factoid Answer Sentence Retrieval", ECIR (European Conference on Information Retrieval)

[46] Wang, Jian & Sun, Jiqing & Lin, Hongfei & Dong, Hualei & Zhang, Shaowu (2016) "Predicting Best Answerers for New Questions: An Approach Leveraging Convolution Neural Networks in Community Question Answering", Chinese National Conference on Social Media Processing, SMP 2016: Social Media Processing pp 29-41

[47] Xu, Huijuan & Saenko, Kate (2016) "Ask, Attend and Answer: Exploring Question-Guided Spatial Attention for Visual Question Answering", European Conference on Computer Vision, ECCV 2016: Computer Vision – ECCV 2016 pp 451-466

[48] Zhou, Guangyou & Zhou, Yin & He, Tingting & Wu, Wensheng (2015) "Learning semantic representation with neural networks for community question answering retrieval", Knowledge-Based Systems, Volume 93, 1 February 2016, Pages 75–83