Institute of Information and Computational Technologies

ISSN : 2788-7677 (Online)
ISSN : 2788-7987 (Print)

ADVANCED TECHNOLOGIES AND COMPUTER SCIENCE

2021
№1

www.atcs.iict.kz

Institute of Information and Computational Technologies,

Advanced Technologies and computer science

№1

Almaty 2021

ISSN: 2788-7677 (Online)

Institute of Information and Computational Technologies,

Advanced Technologies and computer science

28 Shevchenko str., Almaty,Republic of Kazakhstan
7 (727) 272-37-11
atcs@iict.kz

About the Journal

Advance technologies and computer science is a bilingual scientific peer-reviewed, interdisciplinary, electronic journal of open access, including thematic areas:

- Section **"Applied mathematics, computer science and control theory"** includes papers describing modern problems in these areas.
- Section **"Information and telecommunication technologies"** also includes the following topics:
    - Data transmission systems and networks.
    - Internet technologies.
    - Cloud technologies.
    - Parallel computing.
    - Distributed computing.
    - Supercomputer and cluster systems.
    - Big data processing (Big-data).
    - Geographic Information Systems and Technologies.
- In the section **"Artificial intelligence technologies"** in addition to technology, there are works on topics:
    - Intelligent Management Systems.
    - Speech technology and computer linguistics.
    - Pattern Recognition and Image Processing.
    - Bioinformatics and biometric systems.
    - Human-machine interaction.
    - Machine learning.
    - Intelligent Robotic Systems.
- The section **"Information Security and Data Protection"** also covers topics:
    - Software and hardware information protection.
    - Mathematical methods for ensuring information security of complex systems.
- The section **"Modeling and optimization of complex systems and business processes"** may include:
    - Computational mathematics, numerical analysis and programming, mathematical logic.
    - Theory of Statistics.
    - Statistical Methods.

Contents

UDC 004.838.2

IRSTI 20.19.21

COMPARATIVE ANALYSIS OF KNOWLEDGE BASES: CONCEPTNET VS CYC

A. Y. Nuraliyeva[1], S. S. Daukishov[2], R.T. Nassyrova[3]
[1,2]Kazakh-British Technical University, Almaty, Kazakhstan
[3]IT Analyst, Philip Morris Kazakhstan, Almaty, Kazakhstan

[1]nuraliyevaassel@gmail.com, ORCID 0000-0001-6451-3743
[2]suleimen.daukishov@gmail.com, ORCID 0000-0002-3784-4456
[3]nassyrova.r@gmail.com, ORCID 0000-0002-7968-3195

Abstract. Data summarization, question answering, text categorization are some of the tasks knowledge bases are used for. A knowledge base (KB) is a computerized compilation of information about the world. They can be useful for complex tasks and problems in NLP and they comprise both entities and relations. We performed an analytical comparison of two knowledge bases which capture a wide variety of common-sense information - ConceptNet and Cyc.

Manually curated knowledge base Cyc has invested more than 1,000 man-years over the last two decades building a knowledge base that is meant to capture a wide range of common-sense skills. On the other hand, ConceptNet is a free multilingual knowledge base and crowdsourced knowledge initiative that uses a large number of links to connect commonplace items. When looking for common sense logic and answers to questions, ConceptNet is a great place to start.

In this research, two well-known knowledge bases were reviewed - ConceptNet and Cyc - their origin, differences, applications, benefits and disadvantages were covered. The authors hope this paper would be useful for researchers looking for more appropriate knowledge bases for word embedding, word sense disambiguation and natural-language communication.

Key words: knowledge base, natural language processing, data summarization, question answering, text categorization

Intoduction

A lot of research papers [1, 2, 3, 4] are devoted to the construction of Text to Knowledge computing systems. All such projects are united by directions of Deep Learning and technologies - Knowledge Base (KB) and Natural Language Processing (NLP). KB is a repository of information, where it is stored not separately, but in the context of other data units. By extracting data from natural language, we mean a machine that matches nouns with their entities and sentences with their statements. Such a transformation is possible not for any subject area but only where texts are subject to logical discourse and operate on facts (for example, jurisprudence, pharmaceuticals and other hard sciences).

Communication in natural languages is only possible if there is a considerable amount of general knowledge about the world that is shared between the various participants. By creating a knowledge base of general knowledge about the world, as well as specific knowledge in a particular field, they can be useful for a range of complex tasks and problems in NLP.

A knowledge base (KB) is a computerized compilation of information about the world that usually comprises both objects and information, facts about them. Data summarisation [5], named object disambiguation [6], question answering [7, 8], text categorization [9, 10, 11], coreference resolution [12, 13], plagiarism detection [14] are some of the tasks that KBs are used for.

In addition to text processing, structured vocabulary experience with wide coverage is useful for land applications such as geographic information systems and situated robots [15]. The majority of early methods to create KBs were manual. With the advancement of the Web, there are an increasing number of approaches to automatically generate KBs by collecting data from Web corpora. YAGO, DBpedia, Wikidata, NELL, and Google's Information Vault are some of the most well-known approaches. Some of these methods concentrate on Wikipedia, a free online encyclopedia.

Knowledge needed to understand any subject of interest is a key skill for interpreting and deciphering the ever-changing world in the information society. Since much knowledge is conveyed by linguistic communication, whether oral or written, understanding how words are used to convey meaning is critical. Lexical information can be found in a variety of formats, including unstructured terminologies, glossaries, machine-readable dictionaries (e.g. LDOCE [16]), and full computational vocabularies and ontologies (e.g. WordNet [17] and Cyc [18]). Manually developing certain services, on the other hand, is a daunting job. It takes decades to complete and must be done all over again with any new language. Non-English language resources frequently have much less coverage. As a result, analysis in resource-rich languages like English is clearly favored.

Well-known knowledge bases have been developed as a result of the creation of information retrieval approaches focused on the availability of a semantic framework for free text on the internet, such as Google Knowledge Graph [19], NELL [20], YAGO [21], and DBpedia [22]. These databases were created by major companies, such as Wikipedia websites [23], or on a large scale on the internet. Manually curated knowledge bases have also been created, such as the Cyc project [24], which has a small concept space of just 120,000 concepts. The Cyc knowledge base [25], according to Cycorp Inc., is the broadest, deepest, and most comprehensive repository ever built. The Cyc project has spent the last two decades – roughly 1,000 man-years – developing a knowledge base that is intended to capture a broad spectrum of common sense expertise. NELL [21], KnowItAll [26], and ReVerb [27] are examples of automatic approaches. They include many concepts found in scientific texts, but they fall far short of covering many concepts found in scientific articles.

In addition, ConceptNet is a free multilingual KB that uses a rich collection of links to bind everyday entities [28]. ConceptNet is a perfect place to start while searching for common sense reasoning and responses to questions. Finally, BabelNet, which incorporates definitions and relationships from WordNet, the world's biggest semantic lexicon, and Wikipedia, the most frequently cited collaborative encyclopedia.

Today, there are three important challenges in technologies of KB exposed in Cyc and ConceptNet, namely:

• Knowledge is logically contradictory and does not form a theory.

• The prediction for knowledge is poorly defined - the probabilistic assessment of the applicability of knowledge drops dramatically in the process of logical inference.

• Predictions derived from knowledge are statistically ambiguous.

The solution to these problems is discussed, for example, in the work of L. De Raedt and K. Kersting "Probabilistic logic learning". Their solution is closely related to the issue of probabilistic-logical learning, which is the integration of relational or logical representations, probabilistic inference and learning. The author [29] has shown that the solution of the above three problems is possible within the framework of a theory linking two approaches - probabilistic and logical.

Comparative analysis

ConceptNet

ConceptNet [30], a large-scale and publicly available knowledge base comprising millions of common sense statements presented in natural language [31]. ConceptNet was developed as part of the MIT Media Lab's Open Mind Common Sense initiative [32], a crowdsourced knowledge project that began in 1999. ConceptNet's graphical system of linked words/nodes for representing information is particularly useful for textual reasoning over natural language documents. The ConceptNet graph structure, which describes information in patterns of linked word/phrase nodes, is especially useful for textual reasoning over natural language documents. Figure 1 shows an overview of how such information is structured in ConceptNet.

Many practical word processing tasks for real documents are supported by ConceptNet without the need for additional statistical learning, such as thematic list (for example, a news article containing terms like "guns," "demand money," and "get away" may suggest the themes of "robbery" and "crime"), affect-sensitivity (for example, this letter is sad and angry), and analogy (e.g. "scissors", "razor", "nail clipper" and "sword" are perhaps similar to "knife" because they are all "sharp" and can be used to "cut something"),a short description, causal projection, cold grouping of records.

ConceptNet is an information graph that links natural language words and phrases to statements through marked edges. Its information is gathered and revised on a regular basis from a number of outlets, including expert tools, crowdsourcing, and purpose-built games. It is meant to present general language comprehension information in order to enhance natural language applications by helping them to better appreciate the meanings of the words people use.

ConceptNet provides vocabulary and world information from a variety of sources in a variety of languages, as well as acting as a link between knowledge tools. ConceptNet, for example, provides links to URLs defining astronomy in WordNet, Wiktionary, OpenCyc, and DBPedia in addition to its own knowledge of the English language. Embedded word construction is not the only use of ConceptNet, but it is a form of doing so that has strong advantages and is consistent with current distributive semantics study.

For decades, Cyc [19] has been constructing a predicate logic-based common sense information ontology. DBPedia [33] collects material from Wikipedia infoboxes, resulting in a vast number of statistics, mostly for named individuals of Wikipedia articles. Since its content is not publicly accessible, Google Knowledge Database [34] is probably the largest and most general knowledge network.

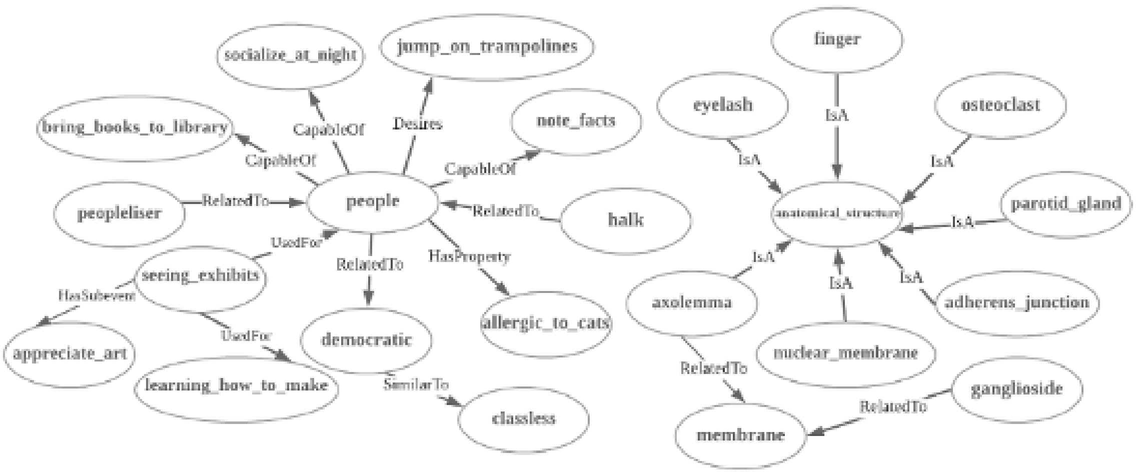

Figure 1 - Fragment of the ConceptNet Knowledge Graph

In contrast to these other tools, ConceptNet's function is to include a broad, free information graph that focuses on the common sense definitions of words as they are used in natural language. Because of the emphasis on words, it is particularly well suited to the concept of representing word meanings as vectors, which is a ConceptNet strength. Word combinations are represented as dense unit vectors of real numbers, with closely related vectors having semantic relationships. This representation is appealing because it portrays context as a continuous space in which relatedness and resemblance can be evaluated.

ConceptNet employs a closed class of relationships, such as IsA, UsedFor, and CapableOf, that are structured to represent relationships independent of the language or source of the words they connect. ConceptNet seeks to match its information services with the 36 partnerships that make up its core package. The intent of these generalised relations is close to that of WordNet relations like hyponym and meronym [35]. Although the edges in ConceptNet are directional, certain connections, such as SimilarTo, are designated as symmetric in ConceptNet 5.5, and the directionality of these edges is irrelevant. Since word embedding will learn from what ConceptNet understands, this knowledge base continues to play an important role in a field that has come to focus on word embedding. As shown by recent findings, ConceptNet may improve the robustness and correlation of word embedding with human judgment.

ConceptNet differs from other common information bases in a number of ways. ConceptNet, in contrast to WordNet [36], which focuses on maintaining lexicographic knowledge and the interaction between words and their context, retains a semantic network system that is intended to collect sensible sense sentences. ConceptNet, in fact, has more relationship forms than WordNet. Cyc [19], a knowledge base that focuses on standardizing common sense for effective logical reasoning, is another comparable knowledge base. ConceptNet, on the other hand, is optimized for inferring from natural language texts and, unlike Cyc, is not a proprietary system.

ConceptNet captures a wide variety of common sense information (as does Cyc), but it does so in a more user-friendly way than higher-order logical notation. ConceptNet has quickly become a useful dataset and resource for different forms of machine learning and NLP over the last decade, owing to these advantages [37, 38, 39].

NLP learning algorithms, especially those focused on embedding terms will benefit greatly from ConceptNet familiarity of graphical construction. ConceptNet can be used to create more effective semantic spaces than distributive semantics.

Various use cases with ConceptNet are given in numerous articles. For example, the paper "Semantic generation mechanism of news images based on Concept-Net" [40] suggests a news image description mechanism based on the ConceptNet knowledge graph. The model authors provide consists of two parts - extracting the image content and rendering with NLP to generate a description of the news image. Another one paper [41] describes enhanced story representation by ConceptNet to predict story endings. The authors propose to improve the representation of stories by simplifying the sentences to key concepts and then modeling the latent relationship between the key ideas within the story. Such enhanced sentence representation, when used with pre-trained language models, makes significant gains in prediction accuracy without using the biased validation data.

Besides, there are several papers about utilizing ConceptNet for sentiment analysis. For example, the paper "KGPChamps at SemEval-2019 Task 3: A deep learning approach to detect emotions in the dialog utterances" [42] describes an approach to solve a task where, given a textual dialogue, the emotion have to be classified according the utterance as one of the following emotion classes: Happy, Sad, Angry or Others. To solve this problem, authors experiment with ConceptNet and word embeddings generated from bi-directional long short-term memory (LSTM) taking input characters. As another example, Ramanathan, Vallikannu, and T. Meyyappan use sentiment analysis to get people feedback about Oman tourism by social media messages in their paper [43]. The authors created their own Oman tourism ontology based on ConceptNet. Entities are identified from the tweets using POS tagger and compared with concepts in the domain specific ontology. After that, the sentiment of the extracted entities are determined by the combined sentiment lexicon approach. Finally, semantic orientations of domain specific features are combined relating to the domain.

In addition, there are some use cases regarding ConceptNet in the educational area. One of them is described in the paper by Su, Ming-Hsiang, Chung-Hsien Wu, and Yi Chang [44] where authors propose an approach to generate follow-up questions based on a populated domain ontology in a conversational interview coaching system. The purpose was to produce the follow-up questions, which are related to the meaning based on the background knowledge in a populated domain ontology. Initially, a convolutional neural network was applied for selecting a key sentence from the user answer. Then the authors used the neural tensor network to model the relationship between the subjects and objects in the resource description framework triple in each predicate from the ConceptNet for domain ontology population. Another work [45] reports on a crowdsourcing experiment conducted with the help of the V-TREL vocabulary trainer (Telegram chatbot) to gather knowledge on word relations to expand ConceptNet. V-TREL offers vocabulary training exercises generated from ConceptNet and collects, assesses the learners' answers to extend ConceptNet with new words.

Moreover, ConceptNet was used in game development. For instance, one of the works introduces the Pokerator, a generator of creative Pokemon names and descriptions, based on user input [46]. The names are generated by mixing words based on syllables or characters according to a bigram language model. A concomitant description is generated by filling a template with ConceptNet answers. Another example is the paper by Lo, Chun Hei, and Luyang Lin, where the purpose is to refine 20Q, a computerized game of twenty yes-or-no questions that asks the player to think of something and the system tries to guess what they are thinking [47]. Authors provide possible research directions, mainly under the formulation of the problem as reinforcement learning. They also investigate methods and potential challenges of incorporating the use of KB for the game.

One more interesting use case is an android application named Talking Diary with a new approach for note classification and scheduling to enable mobile users to automatically organize their daily routine tasks with a single audio note [48]. The application contains three modules: auto audio note classification, auto audio note scheduling, and working hour's calculator. The proposed model of classifier computes similarity score by extracting N-gram weights from ConceptNet to execute classification.

Another research related to robotics proposes a new approach to connect household robot sensor data to Linked Data in order to give robotic agents semantic product information about objects that can be found in their environment, so that the action to be performed with a given object can be inferred [49]. For this, authors use the robot's belief state when recognizing a product and link it to a product ontology that follows Semantic Web standards. Then they use the product class information to fetch further information from ConceptNet that contains action information (e.g. laundry detergent is used for laundering). At last, the action results are mapped to internally known actions of the robotic agent, so that it knows which action can be done with the perceived object.

Cyc

The Cyc Knowledge Base (KB) [19] is the most comprehensive and deepest source of common sense knowledge ever created, with orders of magnitude more content than anyone else. Cyc uses real-world axioms to think about the world and explain the data, but KB is not a database. About 10,000 predicates, millions of sets and definitions, and over 25 million claims are used in Cyc's knowledge base. Cyc will easily prove trillions of bits of used information about the real world when combined with inference engines.

Douglas Lenat began the project at MCC (Microelectronics and Computer Technology Corporation) in July 1984, and the Cyc ontology quickly expanded to about 100,000 terms during the project's first decade, from 1984 to 1994, and included around 1,500,000 words as of 2017. This ontology included 416,000 collections (types, forms, natural species), slightly more than a million individuals comprising 42,500 predicates (relations, characteristics, areas, resources, functions), and roughly a million widely recognised organizations.

A significant number of additional concepts are indirectly used in the Cyc ontology. Cyc's KB of general rules and common sense concepts, which contains these ontological concepts, expanded from 1 million in 1994 to 24.5 million in 2017, which took over 1000 man-years to construct. It's necessary to note that Cyc's ontology engineers tend to hold these statistics as minimal as practicable rather than inflate them, as long as the information base's deductive closure is not impaired.

At different levels of generality, the information in the Cyc KB can be subdivided into loosely clustered, interrelated "blocks" of knowledge as shown in Figure 2. Cyc's interpretation of metaphysics is the broadest block of knowledge, grading down to very basic knowledge of loosely defined domains.Cyc is a long-term artificial intelligence (AI) project that seeks to build a systematic ontology and KB that encompasses fundamental principles and laws on how the world functions. Cyc focuses on tacit information that other AI systems can take for granted in the hopes of capturing common sense knowledge. This is in contrast to evidence that can be discovered on the internet, whether by a search engine or Wikipedia. When faced with new circumstances, Cyc encourages semantic reasoners to execute human reasoning and be less fragile.

The architecture of Cyc's inference engine distinguishes two problems: epistemological and heuristic, that is, what should be in the Cyc KB and how could Cyc efficiently extract arguments hundreds of steps deep in a sea of tens of millions of axioms. The CycL language and well-understood logical inferences could suffice for the former. Cyc used an agent group architecture in the second, in which specialized reasoning modules, each with their own data structure and algorithm, would "raise their hand" if they could work successfully on either of the open sub-problems. There were 20 heuristic level modules (HL) in 1994 [50]. Then there were over 1,050 HL modules in 2017.

The Cyc project's main aim has been to build a large knowledge base containing a stock of formalized context knowledge useful for a range of logic and problem solving activities in various domains since the beginning. Although systems possessing only specialized information of particular domains have produced remarkable outcomes, Cyc's thesis is that these systems are unstable [51] and impossible to apply to modern or unknown problems or problem domains [52]. This is particularly true in areas concerning natural-language communication or responses to questions [53], where the problem domain's breadth is often difficult or impossible to completely describe in advance.

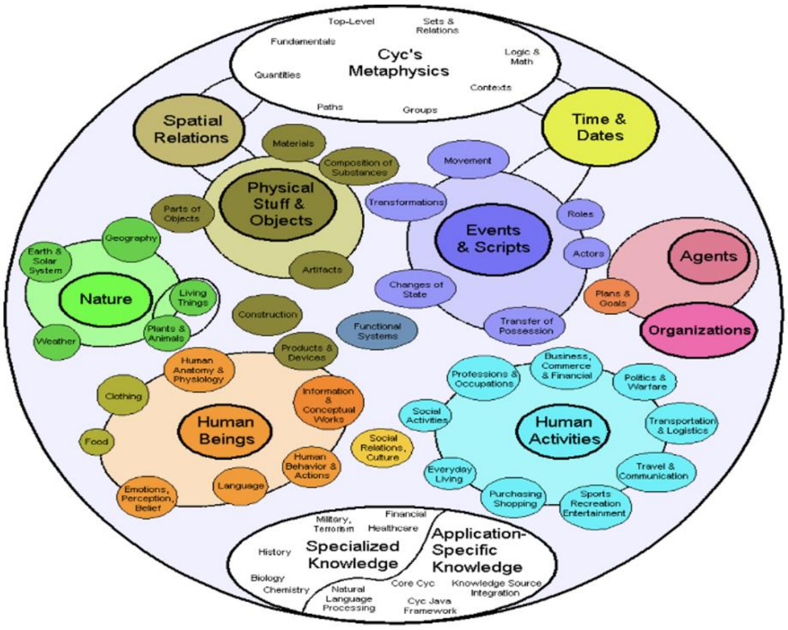

Figure 2 - Cyc knowledge base topic map

This knowledge base is intended to help potential knowledge representation and thinking activities that are unexpected (and also unanticipated). At the time, ontologies or slices of ontology from the Cyc KB were used in a variety of applications, including CycSecure [54], the very early Thesaurus Manager, and the Cyc "Digital Aristotle" program [51].

Applications of the Cyc scheme include:

1. Terrorist identification based on information

The Terrorism Identification Knowledge Base [55] is a Cyc program or framework currently in progress with the goal of preserving all information about "terrorist" organizations, their participants, and those who operate their groups, as well as ideology, founders, guarantors, alliances, programs, locations, capitals, competencies, priorities, events, tactics, and full metaphors for specific terrorist acts. The data is saved as mathematical reasoning files that lead to machine sympathies and cognitive functions.

2. Encyclopedia

An encyclopedia is being created that overlays Cyc keywords on the Wikipedia sheets that are occupied from the sides.

3. Clinical studies: methanalysis

By extending Cyc's ontology and KB about 2%, Cycorp and Cleveland Clinic Foundation (CCF) have built a system to answer clinical researchers' ad hoc queries [56]. The system employs a series of CycL (higher-order logic) pieces of open variables, which are then constrained by medical domain awareness, human-like common sense, grammar, and other factors. There is a way to merge these fragments into a single formal question that is semantically expressive.

4. Healthcare decision support system

The study [57] proposes a knowledge-based system using ontological engineering (by adopting the Cyc method) to assist the creation of a robust foundation for establishing a decision-making support system for the proper diagnosis and management of diseases (e.g. Typhoid Fever, Malaria, Diarrhoea Diseases, Pneumonia, Anaemia) in Sunyani Municipality.

5. Healthcare: effective personalized cancer treatments

The Big C ('C' for Cyc) is a system designed to (semi-)automatically obtain, integrate, and use complex mechanism models related to cancer biology by means of automated reading and a hyper-detailed refinement process resting on Cyc's logical representations and powerful inference mechanisms [58]. Authors' goal is to assist cancer research and treatment with the scale and attention to detail that only computer implementations can provide.

6. Network Security

CycSecureTM is a network risk assessment and network monitoring application that relies on knowledge-based artificial intelligence technologies to improve on traditional network vulnerability assessment [54]. CycSecure integrates public reports of software faults from online databases, data collected automatically from computers on a network and hand-ontologized information about computers and computer networks. This information is stored in Cyc KB and processed by the Cyc inference engine and planner to deliver detailed analyses of the security and vulnerability of networks.

7. Assist in semantic summarization

An abstract description of record data is generated which focuses on the Cyc platform for development. The system's knowledge base and inference engine help it to abstract new ideas that aren't mentioned directly in the document. It makes use of the text's semantic characteristics and syntactic structure. Furthermore, the knowledge base offers subject-matter insight and helps the system to manipulate relationships between ideas in records, which is highly advantageous.

8. Smart AI education

The study [59] offers sixth-grade mathematics learning-by-teaching (LBT) system BELLA built by slightly extending the ontology and KB of Cyc. The "teachable agent" Elle begins with an understanding of the domain content close to the human student's. There is a super-agent (Cyc) which knows the domain content well. BELLA builds up a mental model of the human student by observing them interact with Elle to decide what Elle's current mental model should be to help the user to overcome their current confusions.

Cyc's advantages is that it uses common sense intelligence to dismiss conclusions such as a man becoming pregnant or a 20-year-old having served for 22 years if the inputs suggest so.

To overcome syntactic ambiguities, one must first comprehend the essence of sentences. This enables machines to interpret and comprehend documents written in natural human languages without being caught up in elliptical expressions, overt inconsistencies, or vague sentences.

The Cyc project has been criticized by many AI researchers for lacking a theoretical foundation. However, Lenat contends that while the inference engine can perform deductive reasoning based on database information, it cannot perform induction, that is, it cannot take new data and produce new ideas or relationships. Furthermore, it is 'crystalline' in the sense that assertions are not probabilistically tested. Lenat replied to some of the objections by emphasizing that Cyc is not attempting to create a full artificial general intelligence (AGI) and that it can be quickly integrated into other AI initiatives.

Conclusion

Freebase, WordNet, DBpedia, and Yago are examples of KBs that have been successfully developed and applied to a variety of NLP problems. Machine learning and representational algorithms based on these KBs have advanced year after year in such diverse fields as information graph 'embedding,' question answering, and remote surveillance. ConceptNet and Cyc are the only two cases of common sense logic that we are aware of.

ConceptNet has been described as a semantic network and Cyc has been described as a resource that enables a better comprehension of sentence semantics in a variety of ways. Cyc works on creating a universal schema higher-order logic for expressing sensible semantic statements and can also be used to help reasoning structures [60, 61] that can make more complex logical inferences. New statements are applied to the KB on a regular basis using a mixture of automatic and manual methods.

For decades, Cyc has been constructing a predicate logic-based common sense information ontology. In contrast, ConceptNet's function is to include a broad, free information graph that focuses on the common sense definitions of words as they are used in natural language. Because of the emphasis on words, it is particularly well suited to the concept of representing word meanings as vectors, which is a ConceptNet strength.

Cyc, a knowledge base that focuses on standardizing common sense for effective logical reasoning, is another comparable knowledge base. ConceptNet, on the other hand, is optimized for inferring from natural language texts and, unlike Cyc, is not a proprietary system.

ConceptNet captures a wide variety of common sense information as does Cyc, but it does so in a more user-friendly way than higher-order logical notation. Cyc is created by experts, while ConceptNet is a crowdsourced knowledge project.

NLP learning algorithms, especially those focused on embedding terms, such as word representations in vector space [62], will benefit greatly from ConceptNet familiarity of graphical construction. ConceptNet can be used to create more effective semantic spaces than distributive semantics. On the other hand, Cyc's advantage is that it uses common sense intelligence to dismiss conclusions with no logic.

Regarding known use cases of ConceptNet and Cyc we can conclude that ConceptNet is better represented in the scientific community as there are more relevant articles describing use cases, than we found for Cyc. Besides, ConceptNet is more oriented towards the mass market because there are many use cases in online media, sentiment analysis, game development, and education. Use cases of Cyc are more narrow-profile, such as healthcare, clinical studies, and network security.

References

[1] http://www.opencyc.org [Access date 08042021]

[2] http://www.link.cs.cmu.edu/link/ [Access date 08042021]

[3] http://gate.ac.uk/ [Access date 08042021]

[4] https://github.com/tensorflow/models/tree/master/syntaxnet [Access date 08042021]

[5] Nastase V. Topic-driven multi-document summarization with encyclopedic knowledge and activation spreading. Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing. Waikiki, Honolulu, Hawaii. (2008). 763–772.

[6] Bunescu R., Pasca M. Using encyclopedic knowledge for named entity disambiguation. Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics. Trento, Italy. (2006). 9–16.

[7] Harabagiu S.M., Moldovan D., Pasca M., Mihalcea R., Surdeanu M., Bunescu R., Girju

R., Rus V., Morarescu P. "Boosting knowledge for answer engines". Proceedings of the Ninth Text REtrieval Conference. Gaithersburg, Maryland. (2000). 479–488.

[8] Lita L.V., Hunt W.A., Nyberg E. "Resource analysis for question answering". Companion Volume to the Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics. Barcelona, Spain. (2004). 162–165.

[9] Gabrilovich E., Markovitch S. Wikipedia-based semantic interpretation for natural language processing. Journal of Artificial Intelligence Research (JAIR). 2009. 443–498.

[10] Wang P., Domeniconi C. Building semantic kernels for text classification using Wikipedia. Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Las Vegas, Nevada. 2008. 713–721.

[11] Navigli R., Faralli S., Soroa A., de LacalleO.L., Agirre E. Two birds with one stone: Learning semantic models for Text Categorization and Word Sense Disambiguation. Proceedings of the Twentieth ACM Conference on Information and Knowledge Management. Glasgow, Scotland. (2011). 2317–2320.

[12] Ponzetto S.P., Strube M. Knowledge derived from Wikipedia for computing semantic relatedness. Journal of Artificial Intelligence Research. 2007. 181–212.

[13] Rahman A., Ng V. Narrowing the modeling gap: A cluster-ranking approach to coreference resolution. Journal of Artificial Intelligence Research. 2011. 469–521.

[14] Barron-Cedeno A., Rosso P., Agirre E., Labaka G. Plagiarism detection across distant language pairs. Proceedings of the 23rd International Conference on Computational Linguistics. Beijing, China. (2010). 37–45.

[15] Bateman J. A., Hois J., Ross R., Tenbrink T. A linguistic ontology of space for natural language processing. Artificial Intelligence. 2010. 1027–1071.

[16] Proctor P. Longman Dictionary of Contemporary English. Longman Group, Harlow, UK. 1978.

[17] Miller G.A., Beckwith R., Fellbaum C.D., Gross D., Miller K. WordNet: an online lexical database. International Journal of Lexicography. 1990. 235–244.

[18] Lenat D.B. Cyc: A large-scale investment in knowledge infrastructure. Communications of the ACM. 1995.

[19] Agirre E., Soroa A. Personalizing PageRank for Word Sense Disambiguation. Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics. Athens, Greece. (2009). 33–41.

[20] Von Ahn L., Kedia M., Blum M. Verbosity: A game for collecting common-sense facts. Proceedings of the Conference on Human Factors in Computing Systems. Canada. (2006). 75–78.

[21] Atserias J., Villarejo L., Rigau G., Agirre E., Carroll J., Magnini B., Vossen P. The MEANING multilingual central repository. Proceedings of the 2nd International Global WordNet Conference. Brno, Czech Republic. (2004). 80–210.

[22] Banerjee S., Pedersen T. Extended gloss overlap as a measure of semantic relatedness. Proceedings of the 18th International Joint Conference on Artificial Intelligence. Acapulco, Mexico. (2003). 805–810.

[23] https://www.wikipedia.org [Access date 10042021]

[24] Banko M., Cafarella M.J., Soderland S., Broadhead M., Etzioni O. Open information extraction from the Web. Proceedings of the 20th International Joint Conference on Artificial Intelligence. Hyderabad, India. (2007). 2670–2676.

[25] Chan Y.S., Ng H.T., Zhong Z. NUS-ML: Exploiting parallel texts for word sense disambiguation in the English all-words tasks. Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007). Prague, Czech Republic. (2007). 253–256.

[26] Bizer C., Lehmann J., Kobilarov G., Auer S., Becker C., Cyganiak R., Hellmann S. DBpedia – A crystallization point for the web of data. Journal of Web. Semanti. (2009). 154–165.

[27] Black W., Rodriguez S.E.H., Alkhalifa M., Vossen P., Pease A. Introducing the Arabic WordNet project. Proceedings of the 3rd International Global WordNet Conference. Jeju Island, South Kore. (2006). 295–299.

[28] Harabagiu S.M., Miller G.A., Moldovan D.I. WordNet 2 – a morphologically and semantically enhanced resource. Proceedings of the SIGLEX99 Workshop on Standardizing Lexical Resources. (1999). 1–8.

[29] Vityaev E.E. Knowledge extraction from data. Computer Knowledge Model of cognitive process. Novosibirsk 2006, 293 p.

[30] Hirschberg J., Manning C.D. Advances in natural language processing. Science 2015. 349(6245). 261–266.

[31] Speer R., Havasi C. Conceptnet 5: A large semantic network for relational knowledge. The people’s web meets NLP, theory and applications of natural language processing. 2013. 161–176.

[32] Havasi C., Speer R., Arnold K., Lieberman H., Alonso J., Moeller J. Open mind common sense: Crowdsourcing for common sense. Proceedings of the 2nd AAAI Conference on Collaboratively Built Knowledge Sources and Artificial Intelligence. AAAI Press. (2010). 53.

[33] Lehmann J., Isele R., Jakob M., Jentzsch A., Kontokostas D., Mendes P., Hellmann S., Morsey M., Van Kleef P., Auer S., Bizer C. Dbpedia a large-scale, multilingual knowledge base extracted from wikipedia. Semantic Web Journal. 2014.

[34] GuiLin QI, GAO W.T. Huan. The research advances of knowledge graph. Technology Intelligence Engineering. 2017. 4–25.

[35] Pustejovsky J. The generative lexicon. Computational linguistics. 1991. 17(4). 409–441.

[36] Miller G.A. Wordnet: A lexical database for English. 1995. 38(11). 39–41.

[37] Shi B., Weninger T. Open-world knowledge graph completion. 2017.

[38] Chen M., Tian Y., Chang K.-W., Skiena S., Zaniolo C. Co-training embeddings of knowledge graphs and entity descriptions for cross-lingual entity alignment. 2018.

[39] Lin B.Y., Chen X., Chen J., Ren X. Kagnet: Knowledge-aware graph networks for commonsense reasoning. 2019.

[40] Chen, Zhongmin, Hao Xu, Shuo Gao, and Yongsheng Wang. Semantic generation mechanism of news images based on Concept-Net. In 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA). IEEE (2021). 245-248.

[41] Huang, Shanshan, Kenny Q. Zhu, Qianzi Liao, Libin Shen, and Yinggong Zhao. Enhanced Story Representation by ConceptNet for Predicting Story Endings. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management, (2020). 3277-3280.

[42] Patro, Jasabanta, Nitin Choudhary, Kalpit Chittora, and Animesh Mukherjee. KGPChamps at SemEval-2019 Task 3: A deep learning approach to detect emotions in the dialog utterances. In Proceedings of the 13th International Workshop on Semantic Evaluation, (2019). 241-246.

[43] Ramanathan, Vallikannu, and T. Meyyappan. Twitter text mining for sentiment analysis on people’s feedback about oman tourism. In 2019 4th MEC International Conference on Big Data and Smart City (ICBDSC), IEEE (2019). 1-5.

[44] Su, Ming-Hsiang, Chung-Hsien Wu, and Yi Chang. Follow-Up Question Generation Using Neural Tensor Network-Based Domain Ontology Population in an Interview Coaching System. In INTERSPEECH. 2019. 4185-4189.

[45] Rodosthenous, Christos, Verena Lyding, Federico Sangati, Alexander König, Umair ul Hassan, Lionel Nicolas, Jolita Horbacauskiene, Anisia Katinskaia, and Lavinia Aparaschivei. Using crowdsourced exercises for vocabulary training to expand conceptnet. In Proceedings of The 12th Language Resources and Evaluation Conference, (2020). 307-316.

[46] Geissler, Dominique, Elisa Nguyen, Daphne Theodorakopoulos, and Lorenzo Gatti. Pokérator-Unveil your inner Pokémon. In 11th International Conference on Computational Creativity. (2020).

[47] Lo, Chun Hei, and Luyang Lin. Reinforcement Learning in 20Q Game with Generic Knowledge Bases.

[48] Munir, A. Hani, Abubakar Manzoor, and Utba Aziz. Talking Diary: A Novel Approach

for Automatic Audio Note Categorization and Event Scheduling for Android Application. 2020.

[49] Kümpel, Michaela, Anna de Groot, Ilaria Tiddi, and Michael Beetz. Using linked data to help robots understand product-related actions. In JOWO. 2020.

[50] Lenat D.B., Guha R. V. Building Large Knowledge-Based Systems; Representation and Inference in the Cyc Project (1st ed.). Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1989.

[51] Friedland N.S, Allen, P.G., Witbrock M., Angele J., Staab S., Israel D., Chaudhri,V., Porter B., Barker K., Clark P. Towards a Quantitative, Platform-Independent Analysis of Knowledge Systems, in Proc. of the 9th International Conference on the Principles of Knowledge Representation and Reasoning. Whistler, (2004). 507-515.

[52] Lenat, D. "Cyc: A Large-Scale Investment in Knowledge Infrastructure, CACM 38. 1995. 11.

[53] Hovy E., Hermjakob U., Lin C-Y., Ravichandran D. Using knowledge to facilitate factoid answer pinpointing Full text.. In Proc. of the 19th International Conference on Computational Linguistics. Taipei, Taiwan, (2002). 1.

[54] Shepard B., Matuszek C., Fraser C.B., Wechtenhiser W., Crabbe D., Güngördü Z., Jantos J., Hughes T., Lefkowitz L., Witbrock M., Lenat D., Larson E. A Knowledge-Based Approach to Network Security: Applying Cyc in the Domain of Network Risk Assessment. In Proc. of the 17th Innovative Applications of Artificial Intelligence Conference. Pittsburgh, PA. (2005).

[55] Deaton, Chris, Blake Shepard, Charles Klein, Corrinne Mayans, Brett Summers, Antoine Brusseau, Michael Witbrock, and Doug Lenat. The comprehensive terrorism knowledge base in cyc. In Proceedings of the 2005 International Conference on Intelligence Analysis. (2005).

[56] Lenat, Douglas, Michael Witbrock, David Baxter, Eugene Blackstone, Chris Deaton, Dave Schneider, Jerry Scott, and Blake Shepard. Harnessing Cyc to answer clinical researchers' ad hoc queries. AI Magazine 31, (2010). 3. 13-32.

[57] Appiah S., A. F. Adekoya C. Bapuuroh and C. Akowua-Kwakye. A Knowledge-Base of Prevalent Diseases in Sunyani Municipality, Ethiopia Using Ontological Engineering. J Health Med Informat 2019. 10(338). 2.

[58] Witbrock, Michael, Karen Pittman, Jessica Moszkowicz, Andrew Beck, Dave Schneider, and Douglas Lenat. Cyc and the big c: Reading that produces and uses hypotheses about complex molecular biology mechanisms. In Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence. (2015).

[59] Lenat, Douglas B., and Paula J. Durlach. Reinforcing math knowledge by immersing students in a simulated learning-by-teaching experience. International Journal of Artificial Intelligence in Education. 2014. 24(3). 216-250.

[60] Panton K., Matuszek C., Lenat D., Schneider D., Witbrock M., Siegel N., Shepard B. Common Sense Reasoning – From Cyc to Intelligent Assistant. 2006. 3864. 1–31.

[61] Ramachandran D., Reagan P., Goolsbey K. First-orderized researchcyc: Expressivity and efficiency in a common-sense ontology. 2005.

[62] Mikolov T., Chen K., Corrado G., Dean J. Efficient estimation of word representations in vector space. 2013.

UDC 004.588
IRSTI 20.00.00

TRANSLATES INTO TURKISH LANGUAGES CREATION OF A MODEL OF NEURON MACHINE TRANSLATION

Translates into turkish languages creation of a model of neuron machine translation

Z. Omiotek[1], D.R. Rakhimova[2], A.Zh. Zhunussova[3]
[1]Lublin University of Technology, Lublin, Poland
[2,3]Al-Farabi Kazakh National University, Almaty, Kazakhstan

[1]ORCID ID: https://orcid.org/0000-0002-6614-7799, z.omiotek@pollub.pl,
[2]ORCID ID: https://orcid.org/0000-0003-1427-198X, di.diva@mail.ru
[3]ORCID ID: https://orcid.org/0000-0002-4350-6322, alia_94-22@mail.ru

Abstract. In our work we see that the quality of translation has improved due to the creation of models of translation from Turkish to English and from English to Turkic languages. Turkic-speaking languages are structurally similar. Therefore, studying one of the Turkic languages, you can assemble a corpus for other languages and apply it to the same model. This is done using the OpenNMT model (open neural machine translation). The article shows the morphological, lexical, semantic increase of BlEU (translation index) words and sentences of Turkic languages using OpenNMT. To increase the value of BLEU it is necessary to increase the base in the case. In addition, the work provides a detailed description of the construction of OpenNMT models. Experiments with the Kazakh language, one of the Turkic languages, were conducted and the results were obtained. Words in the Kazakh language taken from the news. The scientific work includes a review of the work of scientists who studied neural machine translation. It is shown that the results of this work outperform the work of other researchers. Having created the neural model of OpenNMT, you will see that the result of the translation is not the same as the online translation from Google, Yandex. OpenNMT also takes less time to read data and saves memory. The results of the experimental study show that the Kazakh-English and English-Kazakh language pairs gave good results in translation.

Keywords: OpenNMT, Turkic languages, neural machine translation.

Introduction

Every year we can see that the quality of machine translation in all languages is improving [1]. These languages include the group of Turkic-speaking languages. Languages that are part of the Turkic-speaking languages: Turkish, Turkmen, Tatar, Kazakh, Kyrgyz, Uzbek, Uyghur, Tuva. However, it is true that Google Translate, Yandex, Prompt machine translations still have many errors in the translation of some complex, negative sentences, idioms, scientific texts.

Assessment of the quality of translating simple sentences using a machine translation Google, Yandex Turkic languages are presented in table 1.

Table 1. Evaluation of the quality of translation of simple sentences by machine translation Google, Yandex for Turkic languages

Source text Turkic languages	Yandex MT	Google MT	Disadvantages
Дустым кичэ матур чалбар кигэн иде (Tatar language)	My friend kiche wore nice trousers	My friend was wearing nice pants last night	Google translator pays attention to punctuation. The Yandex translation of the word "kiche" is incorrect.
Сизни тәбрикләшкә иҗазәт бериң (Uyghur language)	No translation	Let me congratulate you	Uyghur translation is not included in Yandex.

			It is understood as the Tatar language.
Мені жерге қаратпа (Kazakh language)	don't put me on the ground	don't put me on the ground	Directly translated a regular expression
Бу тугрида гап хам булиши мумкин емас (Uzbek language)	It's all in tugrida and can not be found	This is out of the question	The Yandex translation could not translate the word "tugrida" and completely lost the meaning of the sentence
Birsey icmek istiyorum (Turkish language)	I want to drink birsey	I want to drink soething	The Yandex translation replaces the word "something" with the word "birsey"

Table 2. Errors in machine translation of texts of different genres

Genre	The original text is in Russian / Kazakh	Google Translate	Yandex MT
Scientific	The study, published in the journal *Cell*, compared 'brain organoids' -- 3D tissues grown from stem cells which model early brain development -- that were grown from human, gorilla and chimpanzee stem cells. Similar to actual brains, the human brain organoids grew a lot larger than the organoids from other apes. Dr Madeline Lancaster, from the MRC Laboratory of Molecular Biology, who led the study, said: "This provides some of the first insight into what is different about the developing human brain that sets us apart from our closest living relatives, the other great apes. The most striking difference between us and other apes is just how incredibly big our brains are." During the early stages of brain development, neurons are made by stem cells called neural progenitors. These progenitor cells initially have a cylindrical shape	Cell журналында жарияланған зерттеуде мидың ерте дамуын модельдейтін бағаналы жасушалардан өсірілген «мидың органоидтары» - адам, горилла және шимпанзе жасушаларынан өсірілген 3D тіндер салыстырылды. Нақты миға ұқсас, адам миының органоидтары басқа маймылдардың органоидтарынан едәуір өскен. Зерттеуді жүргізген MRC молекулалық биология зертханасының қызметкері, доктор Маделин Ланкастер: «Бұл дамып келе жатқан адам миының бізді ең жақын туыстарымыздан, басқа да маймылдардан ерекшелендіретін ерекшеліктері туралы алғашқы түсінік береді. Біздің басқа маймылдардан айырмашылығы - бұл біздің миымыздың қаншалықты үлкен екендігі ». Мидың дамуының алғашқы	Cell журналында жарияланған зерттеуде "ми органоидтары" - мидың ерте дамуын модельдейтін бағаналы жасушалардан өсетін 3D тіндер-адам, горилла және шимпанзе бағаналы жасушаларынан өсірілген. Нағыз ми сияқты, адам миының органоидтары басқа маймылдардың органоидтеріне қарағанда әлдеқайда көп өсті. Зерттеуді басқарған MRC молекулалық биология зертханасының докторы Мадлин Ланкастер: "бұл бізді ең жақын тірі туыстарымыздан, басқа антропоидты маймылдардан ерекшелейтін дамып келе жатқан адам миынан не ерекшеленетіні туралы алғашқы түсінік береді. Біз бен басқа маймылдар арасындағы ең керемет айырмашылық-бұл біздің миымыз қаншалықты керемет." Мидың дамуының алғашқы кезеңдерінде нейрондарды нейрондық прекурсорлар деп аталатын бағаналы жасушалар жасайды. Бұл прекурсорлық жасушалар бастапқыда цилиндр тәрізді, бұл оларды бірдей

	that makes it easy for them to split into identical daughter cells with the same shape.	кезеңінде нейрондарды жүйке бастаушылары деп аталатын дің жасушалары жасайды. Бұл бастаушы жасушалар бастапқыда цилиндрлік пішінге ие, сондықтан олардың пішіні бірдей еншілес жасушаларға бөлінуін жеңілдетеді.	пішінді бірдей аналық жасушаларға оңай бөлуге мүмкіндік береді.
	They are served as features in our d-gram model replacing the traditional n-gram model.To obtain dependency parse trees, we first applied TreeTagger for tokenization and POS-tagging. Next, we fed the tagged output to the MaltParser that we had trained on the Russian National Corpora.	Олар дәстүрлі n-грамм моделін алмастыратын d-gram моделінің ерекшеліктері ретінде қызмет етеді. Тәуелділікті талдау үшін біз алдымен TreeTagger-ді токенизация және POS-тегтеу үшін қолдандық. Содан кейін біз Ресейдің ұлттық корпорациясында оқыған MaltParser-ге шығарылған өнімдерді жібердік.	Олар біздің d-грамм моделімізде дәстүрлі N-грамм моделін алмастыратын функция ретінде қызмет етеді. Тәуелділіктің синтаксистік талдау ағаштарын алу үшін алдымен treetagger-ті токенизация және POS тегтеу үшін қолдандық. Содан кейін біз Ресейдің ұлттық корпустарында оқыған MaltParser - дің белгіленген нәтижесін тамақтандырдық.
Fiction	Then a lifetime went by. I bought my own apartment in St. Petersburg and was walking around the parental abode collecting my clothes and books while begging my mother for various knickknacks and for old fabrics stored in even older suitcases.	Содан кейін бүкіл өмір өтті. Мен Санкт-Петербургтен өз пәтерімді сатып алдым және ата-анамның үйін аралап жүріп, киімімді және кітаптарымды жинап жүріп, анама түрлі соққылар мен одан да үлкен чемодандарда сақталған ескі маталарды сұрадым.	Содан кейін өмір өтті. Мен өзіме Санкт-Петербургтен пәтер сатып алдым және ата-анамның монастырында жүрдім, киім мен кітаптар жинадым, анамнан одан да ескі чемодандарда сақталған әртүрлі әшекейлер мен ескі маталарды сұрадым.
	What was it like? Maybe something like this: say you have babies, they grow up and turn forty, and you get used to this fact and live with it, and then you rummage in the closet and here he is, your first baby, the way he used to be—eighteen months old, not able to talk yet, sweet-smelling like oatmeal and applesauce, his face puffy from crying, lost but found, waiting for you all	Бұл қандай еді? Мүмкін, осындай нәрсе болуы мүмкін: сіздің балаларыңыз бар деп айтыңыз, олар өсіп, қырыққа келеді, ал сіз осыған үйреніп, онымен өмір сүресіз, сосын сіз шкафта сырласасыз, міне ол сіздің алғашқы балаңыз, ол бұрынғы әдісімен - он сегіз айлық, әлі сөйлей алмайтын, сұлы мен алма тәрізді хош иісті,	Бұл не болды? Мүмкін, сізде балалар бар делік, олар өсіп, қырық жасқа толады, сіз бұл фактке үйреніп, онымен бірге өмір сүресіз, содан кейін шкафқа кіріп кетесіз, міне, ол сіздің алғашқы балаңыз, ол бұрынғыдай-он сегіз айлық, әлі сөйлей алмайтын, сұлы майы мен алма сияқты тәтті иісі бар, көз жасы ісінген, жоғалған, бірақ табылған, сізді осы онжылдықтардың бәрін шкафта күтіп, Сізді шақыра

	these decades behind the storage chest in the closet, unable to call for you—and now you are finally reunited.	жылағаннан беті ісінген, жоғалған, бірақ табылған, сізді осы оншақты жыл бойы шкафтағы сақтау қоймасының артында күтіп тұрған, сізді шақыра алмай - міне, енді сіздер қайта қауыштыңыздар	алмадыңыз—міне, сіз қайта қосылдыңыз.

In the table above, we can see some mistakes in the texts of scientific and literary genres in some complex sentences, in repeated sentences spoken by one author.

Let's take a look at the mistakes made by machine translators using one of the Turkic-speaking languages, a corps of 20,000 lines consisting of Kazakh and English:

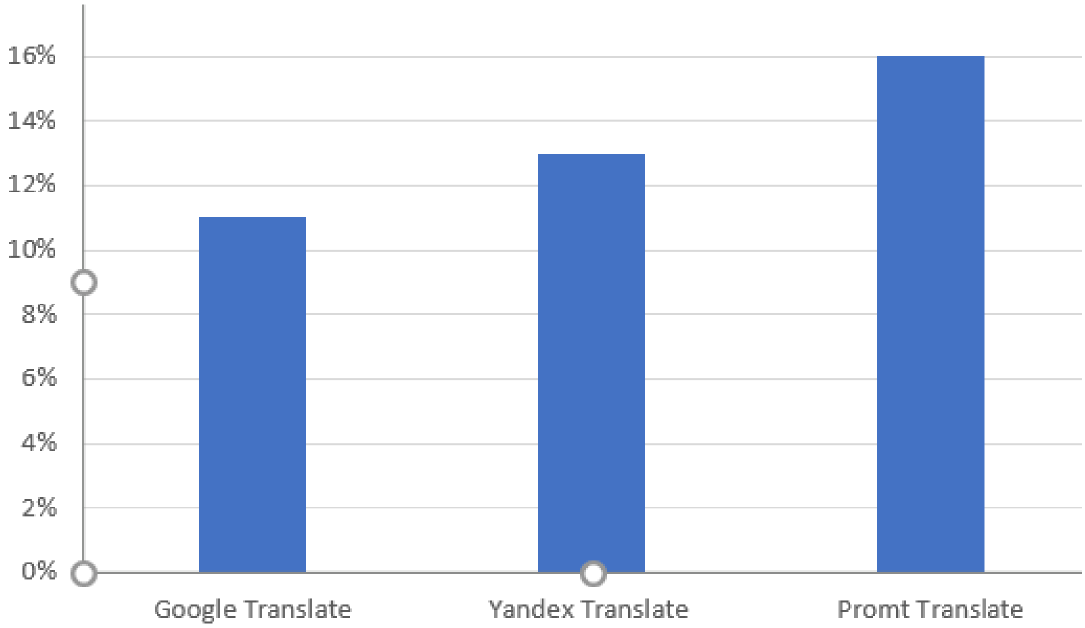

Figure 1- Errors in machine translations

Main part

The article uses the OpenNMT (Open Neural Machine Translation) model of neural machine translation to improve the quality of machine translation.

Section 1. Since the 2000s, Kazakh scholars have been studying the models of machine translations in the translation from Kazakh into other languages. The first research in Kazakhstan was conducted by Tukeyev UA can be seen in the works of [2]. In subsequent works, in 2014-2018, Tukeyev UA, who created a model for statistical machines. students Rakhimova DR [3] and Kartbayev A.Zh. [4] can be noted. Tukeev UA In his works he studied the morphological structure of words in the Kazakh language and created a model of distinguishing suffixes and suffixes [5]. In recent works, Kazakh scientists have used neural machine translation to distinguish morphology. We can see that the translation rate has increased in this work. According to this work, the Kazakh-English parallel corpus was used.

In the present study, data were used, as in [6]. They are produced by the OpenNMT method in a neural machine.

Overview of the structure and model of OpenNMT neural machine translation

OpenNMT is open code created by neurons. Thanks to OpenNMT, data does not take up much memory, machine translation training takes less time. Implementation of training consists of three stages [7]:

1) Data preparation. It consists of words, phrases and sentences compiled from Kazakh and English, respectively, created in parallel in two files.

Data preparation is performed using the code onmt_build_vocab -config toy_en_de.yaml -n_sample 10000. The location of the file must be specified. Write in the YAML configuration file:*data:*

```
data:
    corpus_1:
      path_src: toy-ende/src-train.txt
      path_tgt: toy-ende/tgt-train.txt
  valid:
      path_src: toy-ende/src-val.txt
      path_tgt: toy-ende/tgt-val.txt
```

Add the following information to the YAML configuration file:

Model training:

```
src_vocab: toy-ende/run/example.vocab.src
tgt_vocab: toy-ende/run/example.vocab.tgt
# Train on a single GPU
world_size: 1
gpu_ranks: [0]
# Where to save the checkpoints
save_model: toy-ende/run/model
save_checkpoint_steps: 500
train_steps: 1000
valid_steps: 500
```

In this model we prepare a learning step and a neural model. Then we run this model on code onmt_train -config toy_en_de.yaml.

In translation, it translates 1000 words or sentences from a pre-rendered file. To translate it, write the code onmt_translate -model toy-ende / run / model_step_1000.pt -src toy-ende / src-test.txt -output toy-ende / pred_1000.txt -gpu 0 –verbose.

Section 2. Carry out experimental research on the translation of one of the Turkic languages into Kazakh and show the results

One of the Turkic-speaking languages, Kazakh, was used in the experiment. Kazakh-English and English-Kazakh pairs of neural machine translations were used. A total of 109970 phrases and sentences were compiled and created. The calculation process lasted for 3 days.

Table 3. The result obtained in OpenNMT

Language pair	Speed (tok /s)	BLEU
Kazakh-English	4185	20.56
English-Kazakh	4185	20.05

As a result, it can be seen that the BLEU value is high for Turkic-speaking languages. To give a better result, you need to increase the data in the case.

Conclusion

Summarize the results obtained and note the work that is still under consideration.

By creating an OpenNMT neural model, you can see that getting a translation is no less than online translation from Google, Yandex. However, it can be seen that the translation result can be improved only by increasing the data in the cases. OpenNMT also takes less time to read data and saves memory. As a result of experimental research, it can be seen that the Kazakh-English and English-Kazakh language pairs gave good results in translation. The next study is to obtain translations of other Turkic languages according to this model. For this purpose, a corpus will be assembled for other Turkic-speaking languages.

References

[1] Zhumanov Zh.M., Tukeyev U.A**.** Development of machine translation software logical model (translation from Kazakh into English language). Reports of the Third Congress of the World Mathematical Society of Turkic Countries, (2009). 1. 356-363.

[2] U. Tukeyev , A. Karibayeva & Z h. Zhumanov. Morphological segmentation method for Turkic language neural machine translation. 2020.

[3] Controllable Invariance through Adversarial Feature Learning. Xie, Qizhe, Dai, Zihang, Du, Yulun, Hovy, Eduard, Neubig, Graham. 2017.

[4] Mikel L. Forcado and Ramon P. Neco Recursive Hetero-Associative Memories for Translation. International Work-Conference on Artificial and Natural Neural Networks, IWANN'97 Lanzarote, Canary Islands, Spain, (1997). 453-462.

[5] Tomas Mikolov, Wen-tau Yih, and Geoffrey Zweig. Linguistic Regularities in Continuous Space Word Representations. The Association for Computational Linguistics. In HLTNAACL. 2013. 746–751.

[6] Nal Kalchbrenner, Phil Blunsom. Recurrent Continuous Translation Models. Proceedings of the Conference on Empirical Methods in Natural Language Processing, Seattle, Washington, USA, (2013). 1700–1709.

IRSTI 28.23.15
UDC 004.93

ANALYSIS OF METHODS FOR SOLVING THE PROBLEM OF 3D FACE DETECTION

P. Komada[1], A. Sadykova[2]
[1] Lublin University of Technology, Lublin, Poland
[2]al-Farabi Kazakh National University, Almaty, Kazakhstan
[2]s_akmanat@mail.ru
[1]ORCID ID: https://orcid.org/0000-0002-9032-9285
[2]ORCID ID: https://orcid.org/0000-0002-5512-2297

Abstract. This article considers the problem of face detection in the case of 3D and discusses the current state of the problem, describes the principles of operation of existing 3D detection systems. The fact is that the methods for reconstructing a three-dimensional face model have limitations and drawbacks, which do not allow using them to effectively solve the problem of detecting faces from a video sequence. The methods of restoring the shape from motion, based on matrix factorization, restore the three-dimensional coordinates of only some points of the object, therefore, the problem of interpolating its surface between the restored points of the model arises. Reconstruction of a three-dimensional scene from a pair of images taken at different angles can give an accurate three-dimensional image for almost all points of the original scene, but requires high accuracy of data on the relative position of the cameras. Methods for obtaining a shape from shadows do not allow correctly recovering three-dimensional information about an object in real conditions, when the nature of the illumination is unknown and can be arbitrarily changed.

In research process, the common methods of face recognition in 3D at a given point in time were initially considered. Based on the analysis and research of existing methods, methods for solving a key problem in the field of three-dimensional face recognition - obtaining three-dimensional information about a face are presented.

Keywords: face detection, SFM, PCA, restoration of 3D model of the face.

Introduction

Considered theme must be a grandiose event, most of which have not arrived yet, but are still rising through the present asphalt. Face detection technology is one such occurrence. By the observation, face detection and recognition was mass manufactured in 2017, also now already continued to demonstrate brilliant effectiveness. And terrify whole states to shaking knees.

It is clear that face detection is one of the most promising methods of biometric contactless identification of a person by face. The first face detecting and recognition systems were implemented as programs installed on a computer. Nowadays, facial recognition technology is most often used in video surveillance systems, access control, on a variety of mobile and cloud platforms. The Massachusetts Institute of Technology Journal - MIT Technology Review included facial recognition technology in its 2017 Top 10 Breakthrough Technologies [1].

It is known that one of the main tasks in the field of visual detection is still face detection. A lot of research and development is devoted to this problem, however, the effectiveness of existing face detection systems is still far from human capabilities.

Currently, a relevant and intensively developed direction of research in the field of detection of visual images, especially faces, is the direction associated with obtaining three-dimensional information about an object.

3D detection algorithms [2] use information about the depth and curvature of the surface, unlike 2D detection systems [3] that traditionally use features based on the brightness of image pixels. Consequently, 3D descriptors are more accurate in describing surface features; better suited for describing the properties of the face in the cheeks, forehead and chin; invariant to angle and lighting.

At the moment, there is a task of creating a face detection system based on extracting 3D information from a video stream, which consists in the following: a person sequentially turns his head in three degrees of freedom (tilts forward-backward and left-right, turns left-right) in front of the camera, according to the resulting sequence of frames is built a three-dimensional model of the head and face detection is performed (comparison of the resulting model with the available models in the database).

The main goal of this article and research is to consider modern methods of 3D face detection to solve the problem of developing an effective face detection system.

The current state of the problem of 3D face detection

Currently existing 3D face detection systems use special equipment to reconstruct a three-dimensional face model (sensory systems). Touch 3D detection technologies fall into three categories:

1. Stereo. Two cameras with a known relative position are used to obtain a stereopair of object images; the corresponding points are found on the obtained images and the position of the matched points in three-dimensional space is calculated.

2. Structured light. This approach uses a camera and a light projector: the structural light projects onto the face a special texture, and the camera registers the distortion of this texture on the volumetric object. Using recovery methods shape by texture is calculated the location of points in three-dimensional space.

3. Laser scanning. Laser scanners use light as a source to detect the distance to the scanned object. They measure the reflection time of the laser from the object and receive information about the depth of the points on its surface.

Despite the fact that such technologies give a very high result (detection error within one percent), even with ideal illumination, sensory systems are subject to drawbacks: a frequent case is the appearance of artifacts in the model in the form of "holes" and "protrusions" due to missing data and recovery errors. Another disadvantage of sensor systems is a small depth of field for obtaining the necessary information, for stereo systems - about 0.3 m, for systems with structured lighting - about one meter. Another disadvantage is complex and often expensive equipment.

Nowadays, the following companies that are involved in the development of 3D sensor detection technology can be noted: Geometrix (USA), Genex Technologies (USA), Bioscrypt (Canada), L-1 Identity Solutions (England). In Russia, the Artec Group is working in this direction.

3D face detection systems that do not use additional equipment exist only as experimental developments and do not yet have commercial applications.

Some ways to obtain 3D face information

In order to obtain three-dimensional information about an object, algorithms are used, united in the English-language literature under the name "shape from X" (obtaining a shape using X), where X means a variety of methods [4]. Let us consider those of them that are most promising from the point of view of solving the problem posed in the introduction.

Restoration of the shape by shadows (shape from shading, SFS). The task of restoring the shape of an object by changing the brightness of the image pixels is based on a person's ability to determine the shape of an object using visual information about the nature of light reflection on its surface. This task is a task inverse to the task of visualization (rendering), when the brightness of a point of the simulated scene depends on a number of factors and is calculated according to a given mathematical lighting model [5].

Among the factors affecting the brightness of a point on the surface of an object, the following can be distinguished:

1. Properties and location of light sources.
2. Characteristics of a surface that determine its reflective properties.
3. Orientation of a surface area corresponding to a given point in space.
4. The point of view of the observer

As a mathematical model of the interaction of light and a surface, the Lambert scattering model is usually used, which describes the function of the dependence of the brightness of an image point on the intensity of a single light source, albedo (reflection coefficient) of the surface and the scalar product of the unit normal to the surface and the direction vector to the light source.

Since this mathematical model contains a large number of unknown parameters, in order to reduce the SFS problem to a solvable form, various simplifications are applied, mainly regarding the direction of illumination.

Accordingly, the disadvantages of "shape from shading" algorithms are: the need for a priori knowledge of the law of scattering, too general assumptions about the reflective properties of the surface, leading to incorrect reconstruction of the surface shape. The problem of restoring the shape from filling from the point of view of face detection is most fully covered in [6].

Reconstructing a shape from a stereo pair

Construction of 3D models from a stereopair is traditionally considered as two sequential tasks: stereo comparison and construction of 3D models from a set of points. The task of the algorithm is to obtain data on the distance to objects in the scene, on the basis of which a disparity map is built.

Most of the existing stereo matching algorithms can be divided into two categories of solutions: local and global. Local methods are based on finding feature points and matching them between two frames. Global methods look for a correspondence between images for each pair of pixels, and since there are areas where any texture is missing, smoothness constraints are imposed. Good algorithms look for the displacement map as a piecewise smooth function, with a limited number of discontinuity lines, and take into account that some points are visible in only one image. Generally speaking, local algorithms are less computationally expensive, while global algorithms generate more accurate displacement maps.

Reconstruction of a three-dimensional scene from a stereopair is capable of giving a high-quality result and restoring a three-dimensional image for almost all points of the original image, however, it requires high accuracy of calibration of the stereopair cameras.

With a lack of the spatial structure of an object (the absence of pronounced characteristic points and texture), stereo restoration algorithms find only rough details of the object's shape.

Shape from motion (SFM) restoration

This method of reconstruction of three-dimensional scenes uses the relative movement between the camera and the scene in a sequence of images [7]. As in stereo reconstruction, the SFM problem can be divided into two subtasks: finding a one-to-one correspondence of characteristic points on successive frames and scene reconstruction. But there are also some important differences. The difference between successive frames is much less than between images in a typical stereo pair, since the video is shot at a frequency of several tens of frames per second. Also, unlike stereo, in motion the relative displacement between camera and scene is not necessarily caused by the same 3D transformation.

In terms of matching, the SFM algorithm provides many closely related images (video frames) for analysis, and this is an advantage of this approach. First, tracking techniques can be used here that use the history of movements to predict differences in the next frame. Second, the matching problem can also be viewed as the problem of assessing the apparent motion in an image (optical flow).

As a rule, two types of methods are used to determine correspondences. Differential methods use time derivatives estimates and therefore require a dense sample of sequential images. This method works with every pixel in the image and results in dense measurements. Other methods use a Kalman filter to match and track point characteristics. These methods work with a small number of image points and result in sparse measurements.

In contrast to the problem of finding matches, the problem of reconstruction in this approach is more complicated than that of stereo restoration. Restoring motion and structure frame by frame is more sensitive to noise. This is because the baseline between successive frames is very small.

In the problem of shape restoration from motion, matrix factorization algorithms are used, with the help of which it is possible to restore the position and orientation of cameras, the internal parameters of cameras (focal length), i.e. parameters that are very often unknown. In addition, a large number of frames makes it possible to check the correctness of the matching. Additionally, in a number of cases, you can also obtain estimates of the reconstruction accuracy that correspond to a given scene.

SFM methods based on matrix factorization do not work directly with images, but require as input the coordinates of the characteristic points of the images in pixels and the presence of a marker (number) for each characteristic point, and the same point of the real scene must correspond to the same marker. Thus, the three-dimensional coordinates of only some points of the scene are restored; therefore, the problem arises of interpolating the scene surfaces between the restored points of the model.

3D face detection algorithms

Among the various approaches of 3D detection, three main ones can be distinguished: analysis of the shape of the 3D surface of the face, statistical approaches and the use of a parametric model of the face [8].

Analysis of the shape of a 3D surface. Methods based on the analysis of the shape of a three-dimensional image of a face use local or global characteristics of the surface that describes the face, for example, curvature, line profiles, metrics of distances between two surfaces.

Surface curvature is used to segment the surface of a face into features that can be used to compare surfaces. Another approach is based on 3D face surface descriptors in terms of mean and Gaussian curvature or in terms of distances and angle ratios between feature points of surfaces. Another locally-specific approach is the signature-point approach. The idea of the method is to form a representation-description of a selected point by neighboring points around a given point on the surface. These point signatures are used to compare surfaces.

Global methods use all information about a 3D face image as input to the detection system. For example, a face model is aligned based on its mirror symmetry, and then the face profiles are selected and compared along the alignment plane. It also uses the method of comparing face models based on the maximum and minimum values and the direction of the curvature of the profiles.

Another approach is based on the method of comparing the distances between the detection surfaces. Some methods are based on calculating the metrics of the smallest distances between the surfaces of the models, while others are based on measuring the distance not only between surfaces, but also the texture on the surface. However, a significant limitation of these methods is that the face should not be deformed and its surface is rigid.

The third approach is based on the extraction and analysis of 3D profiles and outlines highlighted on the face.

There are also hybrid methods based on combining local information about the surface in the form of local moments with a global three-dimensional mesh that describes the surface of the entire face. In one of these methods, the value of the function Z (x, y), which describes the depth map of the face in the aligned coordinate system, is decomposed into Fourier components. Decomposition of the function into moments (basis functions) makes it possible to smooth out fine high-frequency “noise on the face” and random outliers. In addition to the Fourier expansion, other basis functions are also used: power series, Legendre polynomials, and Zernike moments.

Statistical methods, in particular the Principal Component Analysis (PCA), were previously widely used in 2D detection. The PCA method was also implemented for 3D detection and was simultaneously extended to a combination of depth and color maps. An alternative for PCA is the linear discriminant analysis method, in which, unlike PCA, one object (a given person) is specified not by one person, but by a set of models (3D faces).

Another statistical method also borrowed from 2D detection is the Hidden Markov Model (HMM) method. The theory of Markov random fields allows one to construct estimates of various spatially variable quantities from images, while imposing certain a priori restrictions on these quantities. Such space-variable quantities can be, for example, offset values in the problem of stereo reconstruction. In the literature on 3D detection, this method is known as pseudo 3D hidden

Markov models and is used, in particular, for recognizing facial expressions.

Using a parametric face model. The key idea of detection by models is based on the so-called parametric 3D models, when the face shape is controlled by a set of parameters (coefficients) of the model. These coefficients describe the 3D shape of the face and can also set the color (texture) on its surface. This method uses one or more facial images as input, mainly photographs taken from the front and profile [9].

The algorithm for solving the problem is built on an iterative principle. As the initial iteration, a certain averaged three-dimensional model of the human head is selected, and its step-by-step improvement is performed. This uses a set of anthropometric points of the face, extracted from the photograph, which is deformed to a given three-dimensional surface. Deformation parameters are calculated during 3D reconstruction using an elastic model. These parameters are then used to recognize this face as a vector.

Conclusion

The above methods of reconstructing a three-dimensional model of a face have limitations and drawbacks that do not allow using them to effectively solve the problem of face detection from a video sequence. The methods of shape recovery from motion, based on matrix factorization, recover the three-dimensional coordinates of only some points of the object, therefore, the problem of interpolating its surface between the reconstructed points of the model arises. Reconstruction of a three-dimensional scene from a pair of images obtained from different angles can give an accurate three-dimensional image for almost all points of the original scene, but it requires high accuracy of data on the relative position of cameras. The methods of obtaining the shape from the shadows are not able to correctly restore the three-dimensional information about the object in real conditions, when the nature of the illumination is unknown and can be arbitrarily changed.

Thus, the most reasonable is the combination of different approaches, proposed in [10]. Algorithms based on matrix factorization are able to provide the necessary information about the three-dimensional coordinates of cameras, their orientation in space and the accuracy with which these values are known. In this case, the methods of stereo matching can give a dense reconstruction of the three-dimensional surface of the object (for each pixel of the image).

In addition, the possibility of determining and comparing the corresponding points of the face in a sequence of video frames based on the construction of a scene illumination model is interesting for further research

References

[1] Top 10 Breakthrough Technologies. URL https://www.technologyreview.com/10-breakthrough-technologies/2017/html (accessed December 12, 2020).

[2] Santos, O. Emotions and personality in adaptive e-learning systems: an affective computing perspective. In Emotions and Personality in Personalized Services. Springer International Publishing. 2016. 263-285.

[3] Wagner, J. and Lingenfelser F. The Social Signal Interpretation Framework (SSI) for Real Time Signal Processing and Recognitions. Proceedings of the INTERSPEECH Florence, Italy (2011). 152-158.

[4] Binali, H. and Potdar, V. Emotion detection state of the art. Proceedings of the CUBE International Information Technology Conference (CUBE '12). ACM, New York, NY, USA (2012). 501-507.

[5] Vezhnevec, V. Graphics and multimedia. URL http://cgm.graphicon.ru/issue2/Paper_vvp2/ (accessed December 15, 2020).

[6] William, A. and Smith, P. Statistical Methods For Facial Shape-from-Shading and Recognition. PhD thesis. University of York. 2007.

[7] Casale S., Russo A., Scebba G., Serrano S. Speech Emotion Classification Using Machine

Learning Algorithms. Proceedings of the IEEE International Conference on Semantic Computing. Santa Clara, CA (2008). 158-165.

[8] Kleinsmith A., Bianchi-Berthouze N. Affective Body Expression Perception and Recognition: A Survey. Proceedings of the IEEE Transactions on Affective Computing. (2013). 4(1). 15-33.

[9] Rosalind, W. Future affective technology for autism and emotion communication. Philosophical Transactions of The Royal Society. B: Biological Sciences. 2009. 364(1535). 3575–3584.

[10] Tao, J and Tan, T. Affective computing: a review. In Proceedings of the First international conference on Affective Computing and Intelligent Interaction (ACII'05), Springer-Verlag, Berlin, Heidelberg (2005). 981-995.

IRSTI 28.23.15
UDC 004.93'1

USING CONVOLUTIONAL NEURAL NETWORKS IN SOLVING PROBLEMS OF IMAGE ANALYSIS AND RECOGNITION

A.Astanayeva[1,2] A.Kozbakova[1]

[1]Institute of Information and Computational Technologies,Almaty, Kazakhstan
[2]Al-Farabi Kazakh National University, Almaty, Kazakhstan

[1]astanayeva@mail.ru, [2] ainur79@mail.ru
[1]ORCID ID: https://orcid.org/0000-0002-6395-5838
[2]ORCID ID: https://orcid.org/0000-0002-5213-4882

Abstract. Convolutional Neural Network (CNN) is a special type of Neural Networks, which has shown exemplary performance on several competitions related to Computer Vision and Image Processing. Some of the exciting application areas of CNN include Image Classification and Segmentation, Object Detection, Video Processing, Natural Language Processing, and Speech Recognition.

The purpose of the work, the results of which are presented in the article, was to research of modern architectures of convolutional neural networks for image recognition. The article discusses such architectures as AlexNet, ZFnet, VGGNet, GoogleNet, ResNet. Based on the results obtained, it was revealed that at the moment the network with the most accurate result is the ResNet convolutional network with an accuracy rate of 3.57%. The advantage of this research is that the given article gives a brief description of the convolutional neural network, and also gives an idea of the modern architectures of convolutional networks, their structure and quality indicators.

Keywords: filter, convolution, neural networks, architecture, deep learning, convolutional neural networks

Introduction

Recognition of visual images is one of the most important components of control systems and information processing, automated systems and decision-making systems. Tasks related to classification and identification of objects, phenomena and signals characterized by the final a set of certain properties and characteristics, arise in such industries as robotics, information retrieval, monitoring and analysis of visual data, artificial intelligence research. With the growth of the computing power of personal computers, as well as the emergence of image databases, it became possible to train deep neural networks. In the task of image recognition, convolutional neural networks (Convolutional Neural Networks) are used. The purpose of the article is to review modern convolutional neural network architectures for the image classification problem.

One of the tasks of machine learning is the task of image classification. To classify an object in an image means to indicate the number to which it belongs recognizable object. For evaluating machine learning algorithms, it is commonly used tagged image databases, e.g. CIFAR-10, ImageNet, PASCAL VOC.[1]

One of the most successful models, considered a recognized leader in the field of image recognition, is convolutional neural network.

Convolutional neural networks (CNNs) are used for optical pattern recognition, image classification, object detection, semantic segmentation, and other tasks. The foundations of the modern SNS architecture were laid in one of the first networks - LeNet-5 by Jan LeKun.

Convolutional neural network structure

The convolution network is a multilayer perceptron (perceptron, English perceptron from

Latin perceptio - perception [2]) - a mathematical or computer model of information perception by the brain, created for recognizing 2D surfaces with high the degree of resistance to scaling, transformations and other types of deformation [3].

Learning to solve such a problem is carried out with reinforcement, with the use of networks of the form, the architecture of which corresponds to the following constraints.

Extraction of features. Each neuron receives an input signal from a local receptive field in the previous layer, extracting its local features. Once a feature is retrieved, its location does not matter, since its location relative to other signs has been approximately established.

Display of signs. Each computational layer of the network consists of many feature maps. Each feature map is shaped like a plane on which all neurons must share the same set of synaptic weights. This form of structural constraint has advantages.

Displacement invariance. Displacement invariance is realized through feature maps using convolution with a small kernel that performs flattening function.

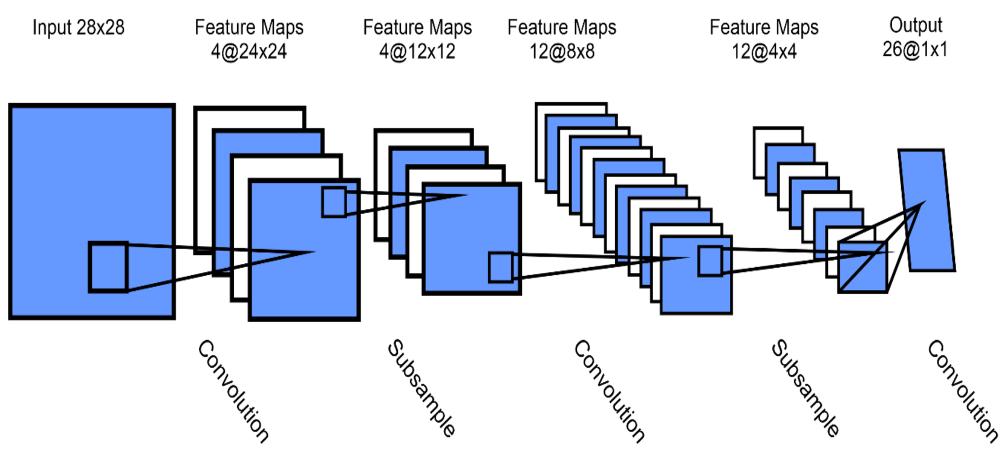

Figure 1 – Convolution network for image processing

Subsampling. Each layer of convolution is followed by a computational layer that implements local averaging and subsampling. By means of local averaging, a reduction in resolution for feature maps is achieved. Such an operation leads to a decrease in the sensitivity of the output signal of the display operator signs, as well as displacement and other forms of transformation.

Figure 1 shows a diagram of a convolutional network consisting of one input, four hidden, and one output layers of neurons. This network was created for image processing, in particular for handwriting recognition. The input layer, consisting of a matrix of 28 × 28 sensor nodes, receives images of various symbols, which are pre-offset to the center and normalized in size. After that, the computational layers alternately implement the convolution and subsample operations.

The first hidden layer is folding. It consists of four feature maps, each of which is a 24 × 24 matrix of neurons. Each neuron has a 5 × 5 sensitivity field.

The second hidden layer performs the local averaging operation as well as subsampling. It also consists of 4 feature maps containing 12 × 12 matrices. Each neuron corresponds to a 2 × 2 receptive field, a sigmoidal activation function, an adjustable threshold, and an adjustable coefficient. The adjustable coefficient and threshold determine the working area of the neuron. For example, with a small coefficient, the neuron operates in a quasilinear mode.

The third hidden layer performs a re-convolution operation. The layer consists of 12 feature maps, each of which is an 8 × 8 matrix of neurons. Each neuron of the third hidden layer can have synaptic connections with different feature maps from the previous hidden layer. The fourth hidden layer performs a second subsampling and repeated local averaging. The layer consists of 12 feature maps, however each feature map contains a matrix of 4 × 4 neurons.

The output layer performs the final convolution step. The layer consists of 26 neurons, each of which corresponds to one of the 26 letters of the Latin alphabet. Each neuron corresponds to a 4 × 4 receptive field [4].

Convolutional neural network architectures for image classification

AlexNet is a convolutional neural network that won the ImageNet LSVRC2012 competition with an error of 15.3%. The network had a partially similar architecture to the LeNet network from J.Lekun, but it was deeper, with more convolutional layers. The network consisted of 11 × 11, 5 × 5, 3 × 3 convolution, max merge, dropout, data increase, function activation of ReLU, SGD (Stochastic Gradient Descent) with an impulse. Function application activation occurred after each convolutional and fully connected layers.

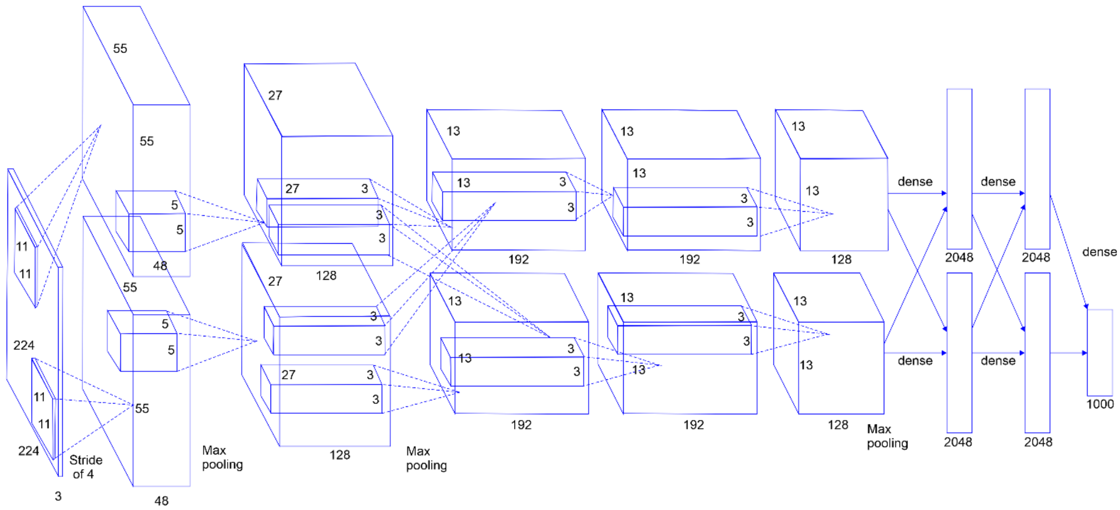

Figure 2 - AlexNet architecture

Shown in Fig. 2. AlexNet architecture, includes 8 layers with weight coefficients, where the first 5 layers are convolutional, the next 3 are fully connected. The output of the last fully connected layer is fed to the softmax activation function, which distributes to the class labels. Further, the network maximizes the goal of the polynomial logistic regression, which is equivalent to maximizing the average over the training cases of the logarithmic probability of the correct label in the prediction distribution.

Using stochastic gradient descent with a learning rate of 0.01, momentum equal to 0.9 and a weight loss of 0.0005. The learning rate is divided by 10 times the accuracy plateau, also decreasing by 3 times during the learning process.

Formulas for updating weight coefficients. Updating the weights (w), where i is the iteration index, v is the momentum variable, and epsilon is the learning rate shown in the diagram. The learning rate was selected equal for all layers, and was also adjusted manually during the entire learning process. The next step was to divide the learning rate by 10, when the number of validation errors stopped decreasing. AlexNet shows the result of top-5 errors - 15.3%, respectively. ZFNet is the winner of ILSVRC 2013 with a top-5 error of 11.2%. The main role in this is played by the adjustment of hyperparameters, namely the size and number of filters, packet size, learning rate, etc. M. Zieler and R. Fergus proposed a system for visualizing kernels, weights, and hidden image representation. The system was named DeconvNet.

The network architecture of ZFNet is almost identical to that of AlexNet. The significant differences between them in architecture are as follows:

- the size of the ZFNet filter and the step in the first convolutional layer in AlexNet is 11 × 11, the step is 4; in ZFNet - 7 × 7, step is 2;
- number of filters in pure convolutional layers of the network (3, 4, 5).

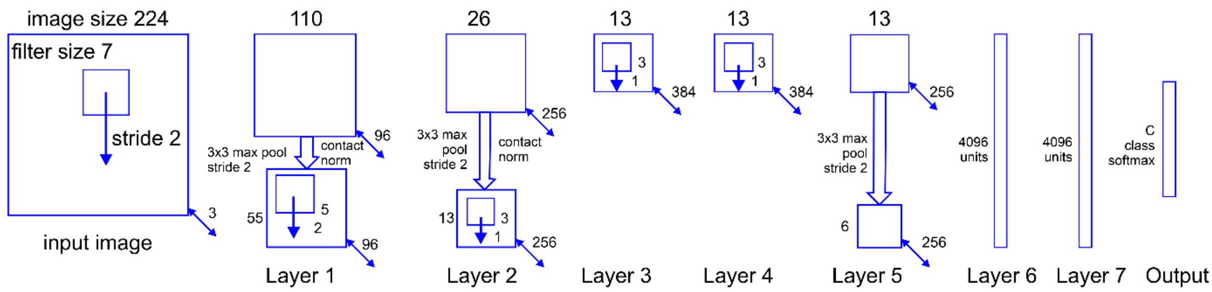

Figure 3 - ZFNet architecture

VGGNet.

In 2014 K. Simonyan and E. Tsisserman from Oxford University proposed a neural network architecture called VGG (Visual Geometry Group). VGG16 is an improved version of AlexNet in which the large filters (size 11 and 5 in the 1st and 2nd convolutional layers) are replaced with several 3x3 filters one after the other.

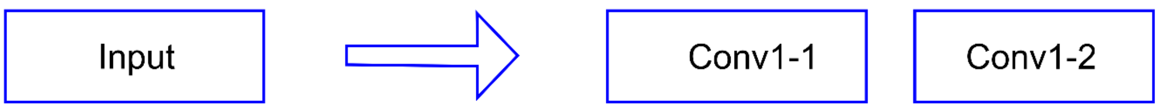

Figure 4 - Architecture of the VGGNet network.

During training, the input to ConvNets (Convolutional networks) is an RGB image of a fixed size of 224 × 224 px. On the in the next step, the image is passed through a 3x3 stack of convolutional layers. In one of the VGGNet network configurations uses 1 × 1 filters that can considered as a linear transformation of the input channels.

The convolution step is fixed at 1 pixel. The spatial padding of the input of the convolutional layer is chosen so that the spatial resolution is preserved after convolution, that is, the padding is 1 for 3x3 convolutional layers. Spatial pooling is done using five max-pooling layers that follow one of the convolutional layers (not all convolutional layers have subsequent maxpooling layers). The max-pooling operation is performed on a 2x2 pixel window with a step of 2.

After the stack of convolutional layers, there are 3 fully connected layers: the first two layers have each 4096 channels, the third layer - 1000 channels (because in the ILSVRC competition it is necessary distribute objects into 1000 categories). The last layer is softmax. All hidden layers are equipped with the ReLU activation function.

The authors have demonstrated that building blocks can be used to achieve specific results in the ImageNet competition. Top-5 errors dropped to 7.3% [5].

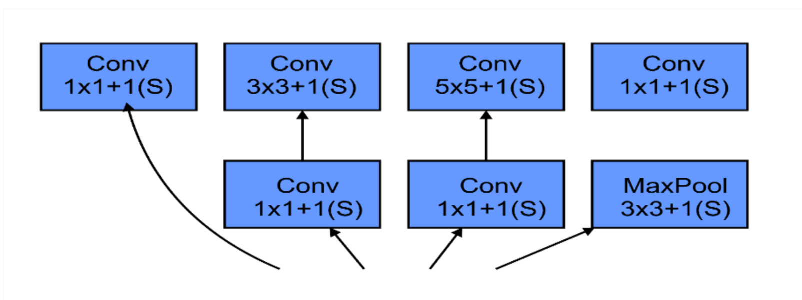

Figure 5 - GoogleNet architecture

A convolutional network from Google (GoogLeNet) known as Inveption-v1 is the winner of the ILSVRC 2014 with a top-5 error of 6.7% [Szegedy et al., 2015].

All convolutions on the network, including those inside Inception modules, use straight-line linear activation. The network has 22 layers when counting only layers with parameters. The total number of layers used to build the network is 100. Moving from fully connected layers to a medium pool improved the accuracy of the top 1 by about 0.6%, but the use of dropout remained necessary even after removing the fully connected layers.

Given the depth of the mesh, being able to propagate gradients back across all layers was an effective challenge. The high performance of smaller networks in this task suggests that the functions created by the layers in the middle of the network must be highly discriminatory. By adding auxiliary classifiers associated with these intermediate levels, discrimination at the lower levels in the classifier was expected. This helped to overcome the vanishing gradient problem by providing regularization. Classifiers in a network take the form of small convolutional networks that are placed on top of the output of Inception modules. During training, their loss is added to the overall weight loss of the net. During inference, these auxiliary networks are discarded. Later control experiments showed that the influence of auxiliary networks is relatively small (about 0.5%) and that only one of them is required to achieve the same effect.

The GoogLeNet architecture uses the Inception module, and the network is built based on modules of this type [1].

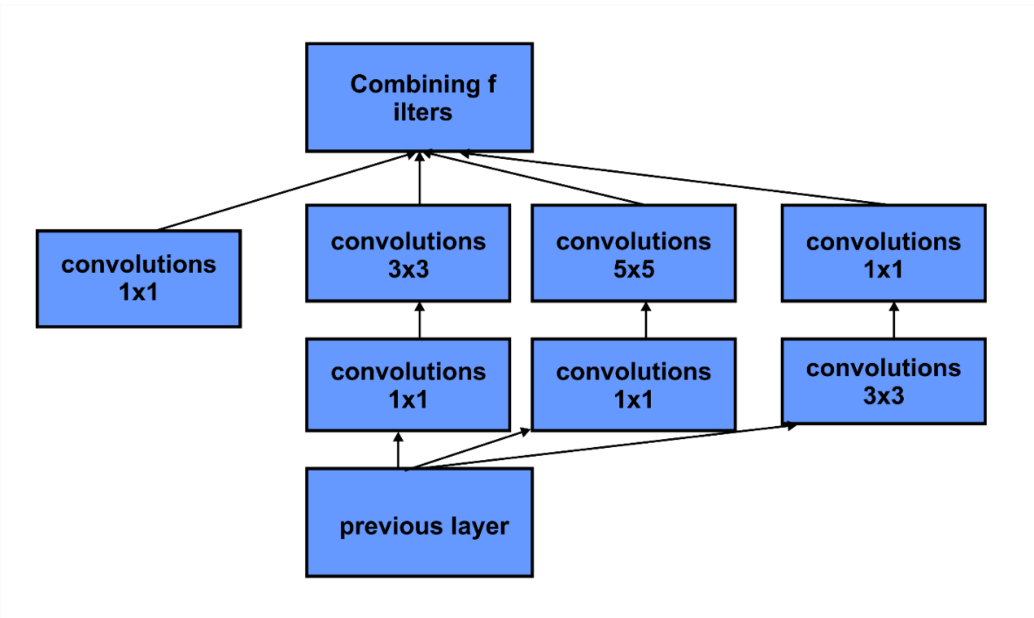

Figure 6 - Module Inception

Inception module. Inception uses multiple (parallel) branches that compute different properties based on the same input and then merge the results together. A 1 × 1 convolution is a way to reduce the dimension of a property map. This type of convolutional layers is presented in the work "Network" in M. Lin's network. As a result, this architecture allows to reduce the number of errors for the top-5 categories by another 0.5% - to the value 6.7%.

Module Inception-v2 and Inception-v3. In the next iteration of the Inception module (Inception-v2 [6]), the layer with a 5 × 5 filter is decomposed into two 3 × 3 layers. The next stage is the use of Batch Normalization [Ioffe, Szegedy, 2015], which allows to increase the learning rate by normalizing the distribution of layer outputs within the network. In the same article, the authors proposed the concept of the Inception-v3 module. In the Inception-v3 module, the principle of filter decomposition is incorporated, namely, the decomposition of an N × N filter with two successive filters 1 × N and N × 1. Also, Inception-v3 uses RMSProp (Adaptive Moving Average Gradient Method) [Hinton, Srivasta, Swersky, 2012], instead of gradient descent, it uses gradient truncation [Pascanu et al., 2013], which is used to improve training stability. A combination of four Inception-v3 modules showed a result in the top-5 category, an error of 3.58% at ILSVRC 2015, Inception-v2 - a top-5 result - 5.60%

ResNet.ResNet is the abbreviated name for the Residual Network (literally - “residualnet").

Convolutional layers have 3 × 3 filters and follow the design rules:

- with the same size of the map of the output objects, the layers have the same number of filters;
- if the size of the feature map is halved, the number of filters, on the contrary, is doubled in order to preserve the complexity of the time for the layer.

Table 1. CNN ResNet characteristics

Layer name	Output size	18-layer	34-layer	50-layer	101-layer	152-layer
Conv1	112×112	7×7, 64, Stride 2				
Conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix}3\times3,64\\3\times3,64\end{bmatrix}\times 2$	$\begin{bmatrix}3\times3,64\\3\times3,64\end{bmatrix}\times 3$	$\begin{bmatrix}1\times1,64\\3\times3,64\\1\times1,256\end{bmatrix}\times 3$	$\begin{bmatrix}3\times3,64\\3\times3,64\\1\times1,256\end{bmatrix}\times 3$	$\begin{bmatrix}3\times3,64\\3\times3,64\\1\times1,256\end{bmatrix}\times 3$
Conv3_x	28×28	$\begin{bmatrix}3\times3,128\\3\times3,128\end{bmatrix}\times 2$	$\begin{bmatrix}3\times3,128\\3\times3,128\end{bmatrix}\times 4$	$\begin{bmatrix}1\times1,128\\3\times3,128\\1\times1,512\end{bmatrix}\times 4$	$\begin{bmatrix}1\times1,128\\3\times3,128\\1\times1,512\end{bmatrix}\times 4$	$\begin{bmatrix}1\times1,128\\3\times3,128\\1\times1,512\end{bmatrix}\times 8$
Conv4_x	14×14	$\begin{bmatrix}3\times3,256\\3\times3,256\end{bmatrix}\times 2$	$\begin{bmatrix}3\times3,256\\3\times3,256\end{bmatrix}\times 6$	$\begin{bmatrix}1\times1,256\\3\times3,256\\1\times1,1024\end{bmatrix}\times 6$	$\begin{bmatrix}1\times1,256\\3\times3,256\\1\times1,1024\end{bmatrix}\times 23$	$\begin{bmatrix}1\times1,256\\3\times3,256\\1\times1,1024\end{bmatrix}\times 36$
Conv5_x	7×7	$\begin{bmatrix}3\times3,512\\3\times3,512\end{bmatrix}\times 2$	$\begin{bmatrix}3\times3,512\\3\times3,512\end{bmatrix}\times 3$	$\begin{bmatrix}1\times1,256\\3\times3,256\\1\times1,2048\end{bmatrix}\times 3$	$\begin{bmatrix}1\times1,256\\3\times3,256\\1\times1,2048\end{bmatrix}\times 3$	$\begin{bmatrix}1\times1,256\\3\times3,256\\1\times1,2048\end{bmatrix}\times 3$
	1×1	average pool, 1000 - dfc, softmax				
FLOPs		$1,8\times10^9$	$3,6\times10^9$	$3,8\times10^9$	$7,6\times10^9$	$11,3\times10^9$

Each ResNet block has two levels of depth (used in small networks such as ResNet 18, 34) or 3 levels (ResNet 50, 101, 152) (Table 1).

50-layer ResNet**:** Each 3-layer block is replaced in the 34-layer network by this 3-layer bottleneck, resulting in a 50-layer ResNet (see Table 1). They use option 2 to increase the dimensions. This model has 3.8 billion FLOPs.

ResNet with 101 and 152 layers: They create ResNet with 101 and 152 layers using more 3-layer blocks (see Table 1). After increasing the depth, 152 layer ResNet (11.3 billion FLOPs) has less complexity than VGG-16 and VGG-19 networks (15.3 / 19.6 billion FLOPs). ResNet - 152 achieves a top 5 result of 3.57%.

Comparison of convolutional neural network models

To assess the performance of convolutional neural network models, indicate the type of error (top-5). The images in the ImageNet database may contain many objects, but only one of them is annotated. The main error criterion is a top-5 error.

The results of comparing the results of various convolutional neural networks are presented in Table 2.

Table 2. Comparison of CNN indicators in image recognition tasks

Neural network	**Top-5**
AlexNet	15,30%
ZF Net	11,20%
VGG Net	7,30%
GoogleLeNet	6,70%
Inception-v2	5,60%
Inception-v3	3,58%
ResNet-152	3,57%

Conclusion

The spread and development of computer vision technologies entails a change in other professional areas of human life. Convolutional neural networks (CNNs) are used in object and face recognition systems, special medical software for image analysis, navigation of cars equipped with autonomous systems, in security systems, and other areas. With the growth of computing power of computers, the advent of image databases, it became possible to train deep neural networks. One of the main tasks of machine learning is the task of image classification. SNS are used for optical recognition of images and objects, object detection, semantic segmentation, etc. In this article, the most common architectures of convolutional neural networks for the task of image recognition, their structure and features were considered. As a result of the analysis of the architectures, it was revealed that the convolutional neural network

ResNet-152 showed the best result in the task of image recognition, with indicator top-5 equal to 3.57%, which indicates a fairly accurate definition of the object.

A feature of the ResNet architecture is that convolutional layers have 3 × 3 filters, and also the fact that a fast connection has been added to the network, which turns the network into its residual version.

References

[1] Sikorsky O.S. Review of convolutional neural networks for the problem of image classification. New information technologies in automated systems. 20. 37–42.

[2] Perceptron. Wikipedia. Last modified 2021. https://en.wikipedia.org/wiki/Perceptron.

[3] Bengio Yoshua, and Yann LeCun. Convolutional Networks for Images, Speech and Time Series, in M.A. Arbib, Ed., The Hand Book of Brain Theory and Neural Networks. Cambridge, MA: MIT Press, 1995.

[4]. Haykin, Simon. Neural Networks: Complete Course, 2nd Edition. Publishing house “Williams”, 2006.

[5] Simonyan Karen, Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. International Conference on Learning Representations, ICLR. (2015). ArXiv preprint arXiv:1409.1556.

[6] Loffe Sergey, Christian Szegedy. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. ArXiv e-Prints. International conference on machine learning (2015). 448-456.

[7] He Kaiming, Zhang Xiangyu, Shaoqing Ren, Sun Jian. Deep Residual Learning for Image Recognition. Proceedings of the IEEE conference on computer vision and pattern recognition (2016). 770-778.

[8] Lee Chen-Yu, Patrick W Gallagher, Zhuowen Tu. Generalizing Pooling Functions in Convolutional Neural Networks: Mixed, Gated, and Tree, Artificial Intelligence and Statistics (2016). 464-472.

INTELLIGENT MODEL FOR CONTROLLING THE TEMPERATURE AND HUMIDITY OF THE BAKING OVEN

I. T.Utepbergenov[1], M.F. Othman[2], SH. D Toibayeva[3], L.S Issabekova[4]
[1]Institute of Information and Computational Technologies,Almaty, Kazakhstan
[2]Putra University, Malaysia

[1]ORCID ID: https://orcid.org/0000-0003-4439-7313
[2]ORCID ID: https://orcid.org/0000-0002-5515-8732
[3]ORCID ID: https://orcid.org/0000-0002-2027-0396
[4]ORCID ID: https://orcid.org/0000-0002-0925-4360

Abstract. The relevance of the study is the need to introduce automated control systems in Kazakhstan. As world experience shows, a promising direction of this task is to improve the quality of food enterprises and reduce production costs. The article is devoted to the problem of automated control of the bread baking process in production. As an example of the development of a fuzzy control model, the process of controlling the temperature and humidity in the bread baking chamber is taken, which is an urgent task. An intelligent control scheme based on a fuzzy controller is proposed.The input and output variables are defined, the formation of decision-making rules by experts is considered, and an example of setting the membership functions of fuzzy variables is given.

Keywords: Automated control, bakery oven, intelligent system, fuzzy model, rule base.

Introduction

The fourth industrial revolution is developing at a new pace in the world. Reacts to optimize production in the real world. A long-term program that provides automation and intellectualization of all industrial production processes from digital product design, creating a digital copy of it to an individualized approach to working with customers. The State program of industrial and innovative development of the Republic of Kazakhstan for 2020-2025 sets 4 tasks for domestic enterprises [1]

The fourth industrial Revolution means greater automation of all production processes and stages. The development of this industrial revolution is an important direction of automation in various types of production. Automation leads to increased productivity, freeing people from the production process, improving the quality of products and fully meeting the needs of society [2,3.].

The introduction of innovative technologies improves the quality of baking and consumer properties of bread, mainly contributes to the competitiveness of enterprises [4]. Information technologies in the bakery industry contribute to changing the relationship between consumption and production, but their interaction requires the exchange of information. [5-7].

A relatively recent approach to data analysis that uses information processing, statistics, and data visualization techniques to find patterns and relationships in large datasets is data mining. Finding their main application in the world of finance and marketing, the methods used have been applied in the processing industry, and more recently in the baking industry [8, 9].

An important factor that can lead to a rethinking of today's achievements and make new theories and practices is data mining, which is a sharp expansion of the capabilities of computer technologies, including hardware implementation of logic and other artificial intelligence tools [10].

Artificial intelligence includes reasoning, natural language processing, and even various algorithms that are used to incorporate intelligence into the system. [11, 12].

In the field of system automation technology development, the potential of various machine learning methods and statistical methods in predicting product quality in production baking processes has been studied [13].

Bread baking and control methods in industrial furnaces

Currently, there are several dozen approaches that are being developed specifically for the study of the management of technical systems. Modern control theory has become a leading science not only in the field of technical, but also in social, economic, environmental, food and other systems [14, 15].

The method of automatic control theory is one of the classical control methods. Disadvantages of this method: when creating a control system, it requires building an accurate mathematical model and taking into account all the constituent factors [16].

With the development of technology, the number of artificial neural networks in various areas of production increases every year. With the help of the neural network, many complex real-world production problems can be solved as a management tool in food technology. [17].

However, they require significant computing power and a relatively longer time for the initial training period [18]

The most optimal method of managing technical objects is the method of artificial intelligence based on fuzzy logic. [19]. It is advisable to use fuzzy logic in cases of a small amount of information about control objects and it is possible to develop effective control systems. [20].

The solution to this problem is the use of fuzzy control methods. [21].

Since fuzzy logic simulates human behavior and decision – making, models and methods based on modeling human thinking and behavior processes are used to describe it, as well as the ability to regulate and change the source data, the measurement range without completely recalculating the regulator. The advantage of fuzzy logic systems over other systems:

- do not require precise mathematical models to perform their functions.
- in complex issues, they can find a quick solution.

Development of a fuzzy control system for temperature and humidity in the baking chamber

Bread baking is a very complex process, consisting of microbiological, colloidal, biochemical, and physical compex processes [22, 23].

The temperature of the test piece changes gradually from the outer layers to the inner layers, and the nature of the processes occurring in it also changes [24]. In the working area, the use of forced heat exchange between the heat carrier and the bread allows you to redistribute the heat flows and ensure a uniform temperature distribution throughout the working area of the baking chamber. To ensure the required heat exchange and humidity in the baking chamber, it is proposed to include a control unit for automatically increasing the heat of the heating element, redistributing the heat flow and turning on the steam humidifier. The required humidity level inside the furnace chamber is provided by the steam supply from the steam humidifier. The temperature in the baking chamber is brought to 240-250 °C (depending on the stages of baking bread). To maintain the set temperature range, temperature sensors are placed in the baking chamber, which are regulated through the control unit (at a temperature difference of ± 10 °C), the inclusion of forced heat exchange fans. To improve the quality of finished products and reduce the baking of bread, you need to adjust the temperature and humidity inside the baking chamber during the baking process. Baking of bread blanks consists of three main stages: moistening, roasting, and baking.

Table 1 shows the technological characteristics of the baking process, taking into account the stages, humidity and temperature.

Table 1. Stages of bread baking in the baking chamber

Allowable limit	Stage 2 Moistening	Stage 2 Roasting	Stage 3 Baking
Baking time	2-3 minute	3 minute	33 minute

Humidity in the chamber	Lower limit	30-35%	11-18%	5%
	Acceptable limit	40-70%	18-24%	5-7%
	High limit	70-80%	24-30%	7-10%
Temperature	Lower limit	100°C	220°C	140°C
	Acceptable limit	110-120°C	240-250°C	150-180°C
	High limit	130°C	260°C	190°C

To create a model for controlling the temperature and humidity in the baking chamber, you need to perform: set the parameters of the fuzzy control model; form the rule bases; set and build the FP; output the solution. [25]. The input and output data are shown in Figure 1.

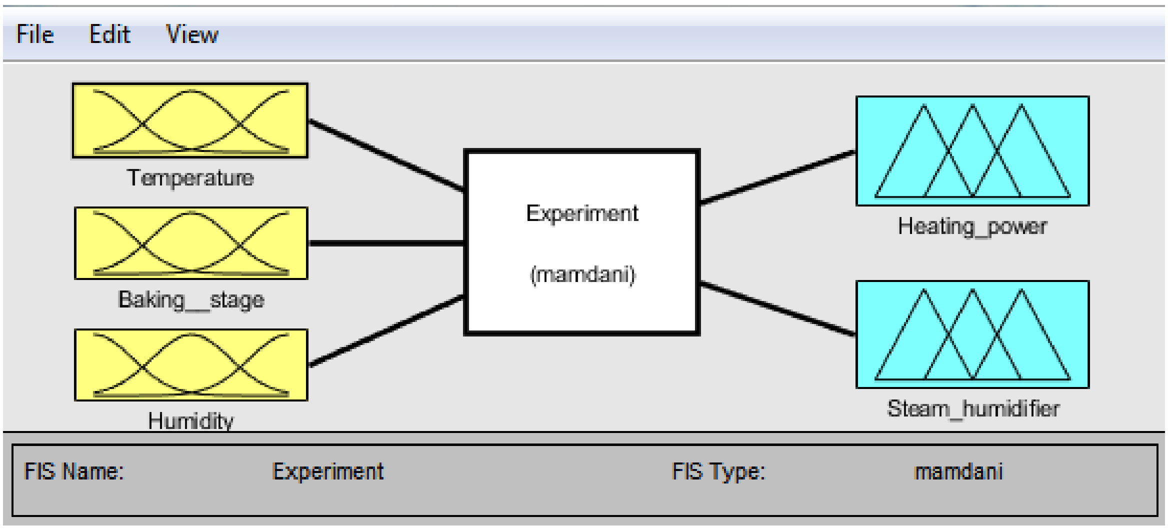

Figure 1-Input and output data

To control the temperature in the baking chamber, we define a set of input fuzzy variables: the temperature in the chamber is 100-300 C, the humidity level is 0-100%, and three stages of baking bread: moistening,roasting, and baking.

We define the set of output linguistic variables by two elements: the regulation of the input power of the heating element of the baking oven and the switching on and off of the steam humidifier.

Based on the qualitative characteristics formulated by experts, you can set a rule base that describes the operation of the control object (OU). "If..., then...", assume to use in the condition of the rule information about the OP-amp, its current state, from the conclusion a control signal is output that brings the OP-amp to the desired state. For example: Rule 1: If in the first stage the temperature in the baking chamber is high and the humidity is low, then the power of the heating element is low and the steam humidifier is turned on. Thus, the rule base of the fuzzy temperature control system in the baking chamber is formed, Table 2.

Table 2. The base of rules of a fuzzy control system

Stage	Temperature	Humidity	Heating power	Steam humidifier
1	Low	Low	Average	Turn on

1	Average	Low	Average	Turn on
1	High	Low	Low	Turn on
1	Low	Average	Average	Switch off
1	Average	Average	Average	Switch off
1	High	Average	Low	Switch off
1	Low	High	Average	Switch off
1	Average	High	Average	Switch off
1	High	High	Low	Switch off
2	Low	Low	High	Switch off
2	Average	Low	Average	Switch off
2	High	Low	Low	Switch off
2	Low	Average	High	Switch off
2	Average	Average	Average	Switch off
2	High	Average	Low	Switch off
2	Low	High	High	Switch off
2	Average	High	Average	Switch off
2	High	High	Low	Switch off
3	Low	Low	High	Turn on
3	Average	Low	Average	Turn on
3	High	Low	Low	Turn on
3	Low	Average	Average	Switch off
3	Average	Average	Average	Switch off
3	High	Average	Low	Turn on
3	Low	High	High	Switch off
3	Average	High	Average	Switch off
3	High	High	Low	Switch off

The fuzzy inference system was developed using Matlab software, will allow the user to continuously adjust three inputs, and the outputs will change according to changes in the input signals:MF

$$\mu(x) = \begin{cases} L\left(\frac{n_1 - x}{\alpha}\right), & x \leq n_1 \\ 1, & n_1 \leq x \leq n_2 \\ R\left(\frac{x - n_2}{\beta}\right), & x \geq n_2 \end{cases} \tag{1}$$

where n is a number called (mode) the mean value of a fuzzy number A $(\mu_A(n))$,

α - is a real positive value of a number or left-handed scatter, β is a right-handed scatter. By increasing the dispersion, we get an even more fuzzy number$A$. Let's write down the fuzzy number L-R:

$$\mu(x) = \frac{1}{1+\left(\frac{x-c}{a}\right)^{2b}} \tag{2}$$

In the Matlab system, the MF is called gbellmf - a generalized bell-shaped MF.

$$\mu(x) = \frac{1}{1+\left(\frac{x-c}{a}\right)^{2b}} \tag{3}$$

For TS "Low" we use z - similar MF zmf
nonlinear approximation

$$\mu(x) = \begin{cases} 1, & x \le a \\ \text{nonlinear approximation}, & a < x < b \\ 0, & x \ge b \end{cases} \quad (4)$$

TS "High" let us describe Smf - s - similar MF

$$\mu(x) = \begin{cases} 1, & x \le a \\ \text{nonlinear approximation}, & a < x < b \\ 0, & x \ge b \end{cases} \quad (5)$$

The output of the "Heating Power" model uses the trapezoidal shape of the trampf accessory for the "Low" and "High" term sets [26]:

$$\begin{cases} 0, & x \le a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1, & b \le x \le c \\ \frac{d-x}{d-c}, & c \le x \le d \\ 0, & d \le x \end{cases} \quad (6)$$

and triangular MF trimpf for the term-set "Average"

$$\begin{cases} 0, & x \le a \\ \frac{x-a}{b-a}, & a \le x \le b \\ \frac{c-x}{c-b}, & b \le x \le c \\ 0, & c \le x \end{cases} \quad (7)$$

For the problem being solved, the number of rules in the rule base is 27.

Table 3. Shows the results of testing the model

№ п/п	Stage	Temperature	Moisture	Heating power	Steam humidifier	Error
1	1	136	81,3	0,175	0,708	0,41
2	1	124	72	0,522	0,711	0,63
3	1	116	35	0,522	0,274	0,53
4	1	97	57	0,846	0,712	1,01
5	1	114	37	0,521	0,273	0,53
6	1	102	37	0,522	0,275	0,57
7	1	90,2	60,8	0,846	0,71	1,03
8	1	27,6	79,9	0,5	0,5	0,93
9	1	119	38,4	0,52	0,49	0,64
10	2	234	16,7	0,5	0,5	0,40
11	2	263	23	0,175	0,276	0,16
12	2	225	18	0,845	0,276	0,46
13	2	256	22	0,35	0,275	0,22
14	2	244	21	0,52	0,275	0,30
15	2	263	20,9	0,175	0,273	0,16
16	2	275	23	0,175	0,273	0,15
17	2	251	18,3	0,521	0,272	0,29
18	2	230	18,3	0,5	0,5	0,40
19	3	179	16,7	0,845	0,267	0,57

20	3	189	12,9	0,523	0,253	0,38
21	3	150	13	0,523	0,253	0,48
22	3	161	12,9	0,523	0,253	0,45
23	3	200	12,5	0,175	0,249	0,20
24	3	143	14	0,517	0,739	0,80
25	3	184	12,3	0,52	0,25	0,39
26	3	1163	11	0,517	0,248	0,07
27	3	182	12	0,519	0,249	0,40

The calculation of the root mean square deviation (RMSD) is made according to the formula

$$\delta = \sqrt{\frac{1}{N}\sum_{i=1}^{m}\left(G_{\text{зад}} - G_{\text{тек}}\right)^2} \tag{8}$$

where n is the number of simulations;
m is the number of hits in the selected interval;
G_{set} – set value of the output variable
G_{cur} is the received value of the output variable.
RMSD does not exceed 5%, which indicates the adequacy of the constructed model.

Conclusion

The standard deviation (RMS) does not exceed 5%, which indicates that the model is built correctly. Thus, it is possible to obtain numerical values of the intelligent control of the cooking chamber using the Mamdani fuzzy logic machine. The correspondence of the constructed model and the influence of the input variables on the output variables were proved using the program for viewing the surface of the blurred model.

References:

[1] Alimkhan A., Makhambayev A., UKAEGBU I. A. The Fourth Industrial Revolution: Towards Energy 4.0 in Kazakhstan. 2019 21st International Conference on Advanced Communication Technology (ICACT). IEEE, (2019). 527-532.

[2] Schwab K. The fourth industrial revolution. Currency, 2017.

[3] Marengo L. Is this time different? A note on automation and labour in the fourth industrial revolution. Journal of Industrial and Business Economics. 2019. 3(46). 323-331.

[4] Mineeva L. N. et al. Prospects of innovative development of bakery products production. Agrarnyy nauchnyy zhurnal. 2017. 9. 87-92.

[5] Waclaw A. , Aloise T. , Lienkamp M. Charging Infrastructure Design for Commercial Company Sites with Battery Electric Vehicles: A Case Study of a Bavarian Bakery. 2020 Fifteenth International Conference on Ecological Vehicles and Renewable Energies (EVER). IEEE, 2020. 1-10.

[6] Baire M. et al. Study and Design of a wireless sensors network for the optimization of bread manufacturing process. 2018 26th Telecommunications Forum (TELFOR). IEEE, 2018. 1-4.

[7] Wang W. et al. Teaching Elementary Linear Algebra for Automatic Control Theory. Chinese Control Conference (CCC). IEEE, 2019. 8950-8953.

[8] Zlatev Z. Plenary Lecture Methods and Tools for Sensors Information Processing. 2020.

[9] Kubicová Ľ., Predanocyová K. Situation in the market of bakery products. International Scientific Days. 2018. 391-406.

[10] Fitriana R., Saragih J., Luthfiana N. Model business intelligence system design of quality products by using data mining in R Bakery Company. IOP Conference Series: Materials Science and Engineering. IOP Publishing, (2017). 1(277). 012005.

[11] Jackson P. C. Introduction to artificial intelligence. Courier Dover Publications, 2019.

[12] Varian H. Artificial intelligence, economics, and industrial organization. National Bureau of Economic Research, 2018. 24839.

[13] Dogan A., Birant D. Machine learning and data mining in manufacturing .Expert Systems with Applications. 2020. 114060.

[14] Pierazzo E. Digital scholarly editing: Theories, models and methods. – Routledge, 2016.

[15] Vassilyev S. N. et al. Intelligent control systems. Procedia Computer Science. (2017). 103. 623-628.

[16] Wang Y. et al. Analysis of energy saving potentials in intelligent manufacturing: A case study of bakery plants. Energy. 2019. 172. 477-486.

[17] Hamamoto A. H. et al. Network anomaly detection system using genetic algorithm and fuzzy logic. Expert Systems with Applications. 2018. 92. 390-402.

[18] Kondakci T., Zhou W. Recent applications of advanced control techniques in food industry. Food and Bioprocess Technology. 2017. 3(10). 522-542.

[19] Rahemi M. R. et al. Evaluation of gamma ray effect on wheat bakery properties in Omid, Roshan and Tabasai cultivars by artificial neural network. 2020.

[20] Bělohlávek R., Dauben J. W., Klir G. J. Fuzzy logic and mathematics: a historical perspective. Oxford University Press, 2017.

[21] Vivek K. et al. Application of fuzzy logic in sensory evaluation of food products: a comprehensive study. Food and Bioprocess Technology. 2020. 1(13). 1-29.

[22] Tsykhanovska I. et al. Research quality of the wheat–rye bread with addition of the polyfunctional food supplement «Magnetofood». EUREKA: Life Sciences. 2017. 5. 45-52.

[23] Balakrishnan V. et al. Thermal flow sensors for harsh environments .Sensors. 2017. 9(17). 2061.

[24] Dobraszczyk B. J., Morgenstern M. P. Rheology and the breadmaking process. Journal of cereal Science. 2003. 3(38). 229-245.

[25] Sabbaghi H., Ziaiifar A. M., Kashaninejad M. Full Research Paper Simulation of fuzzy temperature controller during infrared dry blanching and dehydration of apple slices by intermittent heating method. 2006.

[26] Leonenkov A. V. Nechetkoye modelirovaniye v srede MATLAB i fuzzyTECH. BKHV: Peterburg, 2005.

Responsible for the release: PhD, Shayakhmetova A.S.
Mekebaev A.

Deputy chief editor: PhD, Mamyrbayev O.Zh

Signed in print 03.03.2021
Edition of 50 copies. Format 60x84 1/16. Paper type.
Order No. 4.

Publication of the Institute of Information and Computational Technologies

28 Shevchenko str., Almaty, Republic of Kazakhstan
7 (727) 272-37-11
atcs@iict.kz