ABSTRACT

The diploma work was made on 56 sections, it includes 2 additions and a list of references to the used sources with 19 names. The work contains 19 figures and 1 table.

**Relevance of the theme.** Today's modern medical companies need to have a published archive of new videos for further effective processing, obtaining analytical data, successful communication with customers. This very task of classifying new reports emerged on the TV channel Ukraine. Novelty channels perform classification and processing of news stories manually or with the help of contractors, but this task can be accomplished due to an individual intellect. The problem is relevant, because there is no such system in use in any of the Ukrainian media companies.

**Objective and tasks of the research.** The aim of this thesis is to create a mathematical and software for processing and classifying the archive of new video reports.

* When developing the appropriate software, the following tasks had to be fulfilled:
* Carrying out a comparative analysis of the existing methods of nlp and cv;
* Choice and adaptation of the existing method for solving the nlp and cv tasks;
* Development of the software on the basis of the chosen mathematical method;
* Testing of the developed automated system.

The implemented system must satisfy the following requirements:

* Identify the day of the calendar year from the video file name;
* To identify the start-end time of each individual scene;
* Identify story timing;
* Identify the title of the story;
* Designate the text of the story;
* Identify the type (news, announcement, or advertisement);
* Identify the journalists and guests;
* To identify the anchors;
* Identify tags;
* Classify the story as positive or negative;
* Add the channel and the name of the new program to the result table;
* Have high efficiency indicators of detecting start-end time of the story (accuracy should be up to 1 second);
* Present the table of data processing of the new release in the form of a Microsoft Excel file;

**Research methods.** This work analyses existing approaches to fulfillment of the set task. Methods of solving the problem are related to such modern areas of information processing as computer vision and natural language processing.

**Characteristics of the developed system.** This work analyzes existing approaches and creates an automated system for detection of the system was designed and implemented a software system that allows the processing and classification of the archive of new video reports.

**Results of system tests.** With the help of created TOR was processed more than 80 years of new releases, for each group of results analyzed their accuracy by comparing them with the data obtained manually, the accuracy of the obtained results was set at 85-100% for each category.

**Practical Use.** The developed software was put into operation for the TV channel "Ukraine". The next step is to adjust the algorithm for other news channels such as Ukrainian24, Inter, STB, ICTV, etc.

**Keywords:** media stories, video stream processing, news, natural language processing, computer vision.