A Systematic Mapping Study on Artificial Intelligence Tools Used in Video Editing

Igor Bieda [1†] **and Taras Panchenko** [2††],

tp@infosoft.ua

Department of Theory and Technology of Programming

Faculty of Computer Science and Cybernetics

Taras Shevchenko National University of Kyiv

Akademika Glushkova Ave., 4-D, Kyiv, Ukraine, 03680

Summary

From the past two eras, artificial intelligence has gained the attention of researchers of all research areas. Video editing is a task in the list that starts leveraging the blessing of Artificial Intelligence (AI). Since AI promises to make technology better use of human life although video editing technology is not new yet it is adopting new technologies like AI to become more powerful and sophisticated for video editors as well as users. Like other technologies, video editing will also be facilitated by the majestic power of AI in near future. There has been a lot of research that uses AI in video editing, yet there is no comprehensive literature review that systematically finds all of this work on one page so that new researchers can find research gaps in that area. In this research we conducted a statically approach called, systematic mapping study, to find answers to pre-proposed research questions. The aim and objective of this research are to find research gaps in our topic under discussion.

Key words:

video editing, artificial intelligence, systematic mapping study.

1. Introduction

From all content available on the internet, video is the most appealing and popular stuff. More than 75% of content on the internet is based on videos capturing or making videos has been an easy task with help of mobile phones and social media tools [1]. Even then editing these videos is still a tedious task. Video editing suffers from many challenges such as they need to handle video as well as video simultaneously, and handling video frame by frame are the challenges that make this task very difficult. Numerous attempts have been made to make this job easier.

Video editing is an art of joining and cutting different pieces gathered from diverse sources [2]. Video editing tools are basically no more than computer programs that perform video editing in a sophisticated and easy-to-handle manner. Artificial Intelligence (AI) has also played a vital role to make video editing easy and time-efficient. Rapid growth of AI techniques makes tasks of the real world easier than ever. Branches of AI, like machine learning, reinforcement learning, deep learning, and computer vision, etc. brought revolution in scientific innovations. Intelligent video editing tools are playing their role in facilitating video editing. The pivotal feature of an intelligent video editing tool is to enable the manipulation of video from high level of abstraction. Intelligent video editing tools rely on dialogues rather than frames. The first intelligent video editing tool used was Silver [3]. The Silverworks on automatic selection of video clips based on predefined criteria. Rough-cut is another example of a video editing tool. It edits video computationally, by user inputs for dialogues [4]. Both are text-based video editing tools. They use subtitling techniques in interview videos. QuickCut is a toll that was developed particularly for narrative videos while video digest was a worthy contribution in lecture videos [5]. Lecture videos are a branch of videos that specifically deals with education material such as tutorials etc [6].

In recent years hype has been created to make complete automated video editing. Mashups and video summaries are cutting edge technologies gain tremendous fame in the past half-decade. These involve particular requests made by simple algorithms hence no manpower or extra intelligence is required for the technologies. To make a complete automatic video editing tools take attention of many researchers [7]. New AI technologies, like machine learning, reinforcement learning, computer vision, deep learning, and object tracking, etc. play a great role to make video editing tasks more easy and efficient. Still, horizons to go to make video editing fully automated. Rationalization of the video editing process has still a big challenge for their users. So far, editing is done by examining video clips frame by frame. This is a time-consuming and tiresome task. Numerous solutions for this problem have been presented so far. Soe presents a survey that base on interviews on video editing experts. In this research he presented users and video interaction of AI technology. From our best knowledge there has not been so far any statically arranged survey that presents broad range of AI tools used in video editing research. Hence, well comprehensive review is much needed to bring all solutions on one page, to let research choose their specific area [7]. Our contribution in this research is to make a literature review that will help the new researchers in finding gaps in that particular research area. We conducted

a Systematic Mapping Study (SMS) in this research. SMS is a kind of literature review that statically finds answers to particular pre-defined research questions.

The purpose of SMS is to assist the new researchers in choosing their specific research domain from the big canvas. This research will not focus on video editing commercial software, rather it attempts to cover research work in the aforementioned field. Furthermore, the categories we developed in section IV are not depicting commercially available video editing techniques; rather we make this categorization according to our search. Similarly, AI tools mean by techniques of AI, rather than pre-developed tools. In other words, our research is focusing research area of AI and video editing and finding research gaps in those areas.

Reset of our paper presents background/preliminaries in section II, research methodology in section III, the result of an experiment in section III, and discussion in section IV along with conclusion in future work at the end.

1.1. Preliminaries

In this section, we will discuss basic concepts of video editing, not particularly for this research, but for initial domain knowledge. Since our research is targeting two major technologies, video editing and artificial intelligence, and its techniques, we are going to discuss preliminary domain knowledge of both domains.

1.2. Video editing terms

The following terms one must know about video editing, but this is from a technical point of view, not for this particular research.

- Aspect ratio: aspect ratio deals with the ratio of width and height of the video. Examples of aspect ratios are the following 4:3, 16:9, 1.85:1, where 4,16 and 1.85 are height and 3,9 and 1 are width of video.
- B-Role: b-roll deals with the background seen. This term is used for the factors that help scenes like weddings, sports, news, etc.
- Bit Rate: bit rate, also called data rate deals with the amount of data that is used by video in each second.
- Close-ups: close-ups are the term, deals with shots that are being tightly the frame. For instance, the close of a person will cover the whole face of that person.
- Color temperature: this property of video editing deals with light that is visible in any shot. It is in the range from cool to hot. As the color is hot it seems more visible light in the shot and the light colors pretend less light.
- Compositing: compositing is the most commonly used term in video editing. It deals with merging two or more frames for editing the video.

compositing is mostly used in digital image processing.

- Compression: as its name suggested compression is the process of reducing the size of a video. As the size is reduced as quicker it will be uploading and downloading. Compression techniques that are used in the video are not much different from compression techniques used for other data items.
- Crop factor: crop factor is a ratio between full areas of the screen to the particular area that needs to be sneered. Mostly this ratio ranges from 1.3 to 2.0.
- Foley: the term foley is mostly used for multimedia video editing particularly movies, dramas, etc. it deals with a sound that is being mixed after the video is captured. For instance, the voice of chewing or nature voice cannot be captured in the film, they just reproduce in the studio.

1.3. Types of video editing

To take an insight into video editing, one must know its types. According to our best knowledge video editing can be said to form one of the following types. Linear video editing: this is a type of video editing that uses videotape. Since is linear nature of editing is called its name linear video editing. In linear video editing, numerous video clips from different resources are gathered in a single video [8].

- Non-linear editing: this type of editing called non-linear editing system, also called NLE, edits video from specific computer software [9]. Numerous NLEs are commercially available such as DaVinci Resolve, Adobe Premiere, and Avid Media Composer, etc.
- Offline editing: this is the kind of video editing that deals with copying raw footage from its original source and after editing it are again embedded into that source, hence making a new video and the original video remains unchanged [10].
- Online editing: online editing is a successor processor of offline editing. The video that has been edited offline, reassembled for full resolution editing. This is the final stage of video editing [11].
- Cloud-based editing: cloud-based video editing uses the internet to edit remotely [12]. This is mostly used for applications that are time-critical such as sports coverage etc. in cloud-based video editing low-resolution copies is being used.
- Vision mixing: this type of video editing is used in live television shows. Vision mixers help in

cutting live feeds that come across from different sources such as cameras in real-time.

1.4. Comparisons of video editing software

In this section, we will make a comparison of different commercially available video editing software in **Table 1**.

Table 1: Comparison of commercially available video editing software[1]

Program	Platform	Year of the first release	Year of the latest release	License
Adobe Premiere Elements.	Windows, Mac	2004	2020	Trialware
Adobe Premiere Pro.	Windows, Mac	1991	2020 (Adobe CC 2020)	Commercial
Autodesk Flame	Linux (CentOS), Mac	1993	2020	Commercial
Autodesk Smoke	Mac	?	2020	Commercial
Avid Media Composer	Windows, Mac	1989	2020	Commercial

2. Research Method

In this section, we will briefly discuss all steps of our methodology. The purpose of this methodology is to discover answers to research questions. Hence in this section, first of all, we present our research questions, afterward, we will discuss steps of our review that include, choose data sources, determining the research query for different data sources, initial findings, and inclusion and exclusion criteria and at the end, processing or classification of our collected data. We taxonomies research domain of video editing according to our best knowledge.

2.1 Research Questions

For comprehensive findings of the use of AI tools in video editing, we identify the following research questions.

1. How has the trend of using AI tools and techniques in video editing grow historically?
2. Which AI techniques are more frequently used in particular video editing categories?
3. In which venue, journal, conference, symposium video editing papers have been published?
4. Which AI tool and technique can be used in video editing according to the research gap?

3. Result and discussion

3.1. Data search

- We choose following six data sources for conducting our search using particular keyword in each data source format in **Table 2**.

- IEEE Xplorer
- ACM digital library
- Scopus
- Sage
- Nature
- Web of science

Table 2 : Initial result from selected data sources

Data source	Search string	Paper found
IEEE	("Abstract": video editing*) AND ("Abstract": artificial intelligence OR AI OR tools OR techniques)	423
ACM	[All: "video editing"] AND [All: "AI" "tools" "techniques"]	164
Scopus	TITLE(video editing) AND TITLE-ABS-KEY (artificial intelligence OR AI OR tools OR technique)	113
science direct	Title, abstract or author-specified keywords: (video editing) AND (OR artificial intelligence OR AI OR tools techniques)	154
Nature	(Abstract:("AI")) AND (Abstract:("video editing"))	59
Sage	[Title video editing] AND [[artificial intelligence] OR [All tools] OR [All technique]] OR [All AI tools]]	106

---

[1]https://en.wikipedia.org/wiki/Comparison_of_video_editing_software.

Table 2 presets our initial finding from different data sources. We choose date range from 1970 to 2021 to that we can investigate all historical data. These papers are further filtered by inclusion/exclusion criteria.

3.2. Inclusion/exclusion criteria

An inclusion-exclusion criterion is very important in a systematic mapping study. It defines how to refine our initial findings up to the next level of study. Each paper was evaluated on particular inclusion and exclusion criteria.

Inclusion criteria:
- paper that published in any conference or general or symposium;
- paper that is focused in the domain of video editing;
- paper used AI terminology like machine learning, computer vision, deep learning, object tracking and reinforcement learning etc.

Exclusion criteria:
- book chapter;
- paper that does not particularly focus on video editing;
- paper that only presents domain knowledge or guideline of video editing;
- duplicate search of the same study in different data sources;
- papers, written in a language other than English;
- papers that answer at least one or more research question.

3.3. Data extraction

In the data extraction step, we look at all the included papers and gather as much information to Answer the research questions as possible, as well as looking at data that is not related to the research Question but still interesting for the systematic mapping study. In this thesis, we combined the data extraction step with the prerequisite, the quality assessment. The quality assessment has the function to check through all included papers and determine if they hold valuable information or if they do not mention anything to answer the research questions. The data that was gathered is shown in the data extraction form of **Table 3.**

Table 3: All field of data

Data Item	Value	RQ
Title	Title of research article.	
Abstract	Abstract of paper	

Relevance Code	TRUE/False (After applying exclusion criteria it contain only TRUE value)	
Key Words	Keywords	
Authors	Name of all authors	
Year of publication	Year of publication	RQ1
Venue of publication	Name of conference/symposium or general	RQ2
Venue type	Conference / symposium / workshop / general	RQ2
Video editing sub domain	Name of video editing sub domain (mentioned in table 4)	RQ3
AI techniques	Name of AI techniques (mentioned in table 5)	RQ4
Data source	Name of digital library	

Based on our extracted data, we taxonomies research the field of video editing in categories mentioned in **Table 4** and AI techniques used in video editing in **Table 5**. We design this taxonomy based on our best knowledge. Since our focus in this research is particularly the use of AI tools and techniques used in video editing, we did not take an insight into technical aspects of video editing. Our categories are based on 'key words' used in the papers, filtered after inclusion/exclusion criteria.

Table 4: Video editing research taxonomy

Title	Description
Automatic video editing	It is process of choosing desired footages from the video automatically and making a new video by coherent fotages [1]
AR/VR Content creation	It deals with creation of augmented reality and virtual reality content in video [13]
Compressed domain	Video those are initially in compressed domain [14]
Deepfake	Deep fake is a promotion of deep learning. It deals with replacement of image of person to another person [15]
Home video editing	This is category of video editing software that allows home user to edit their videos [16]
Geodestic image	These are n dimension images [17]
Multimedia	Deals with video editing of

| communication | multimedia applications like movies etc. |
| Gradient domain | Is a type of digital image processing that deals with neighboring pixels rather than value of picture directly [18] |

In **Table 5** we highlight particular AI techniques used in our extracted papers. These AI techniques or sub-domains are categorized specifically for our under discussion search, hence we do not claim the definition and categorization as generic for all users.

Table 5: AI techniques used for video editing

Title	Description
Machine Learning	Machine learning is an emerging branch of AI. It deals with training algorithms so that they can improve them self
Deep Learning	Deep learning can be categorized as a special specific branch of Machine learning that works like human brain
Computer vision	Computer vision is branch of AI that particularly deals with audio and video algorithms etc
Digital image processing	It deals with processing digital image using algorithms on digital computers [19]
Reinforcement learning	Reinforcement learning as its name suggested, a specified form of machine learning that deals with reward good behavior and punishes bad behavior of algorithm.
Object tracking	Object tracking is a special form of deep learning. It takes an object at its initial level and then detects its presence in the frame in any later time.

For the rest of this section, we will discuss answer to all research questions.

- RQ1: out first and most important question is related to the increasing trend of using AI toll and technique particularly in the video editing field. **Fig. 1** shows the yearly trend of using AI-based techniques in video editing. This graph shows that the use of AI techniques is increasing in this domain year by year. The insight of this graph shows particular AI areas such as computer vision, machine learning, deep learning, etc. Hence answer to our first RQ is very positive and encouraging for the new researcher. Computer vision and object tracking are the most commonly used AI techniques in video editing research. As this chart shows increasing trend of using AI tools in video editing, year 2021, is the highest publishing year in this research area. One other point to notice is from recent years like 2017 all AI techniques is widely being adopted in video editing research. Hence, this is encouraging sign for researcher of video editing area to use research related to AI in their field.
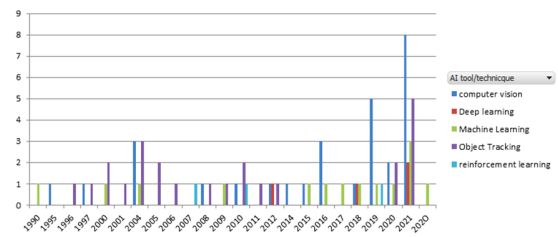

Fig 1. Trend of using AI toll in video editing

- RQ2: our second research questions deals more specifically with AI techniques used in branch of video editing.
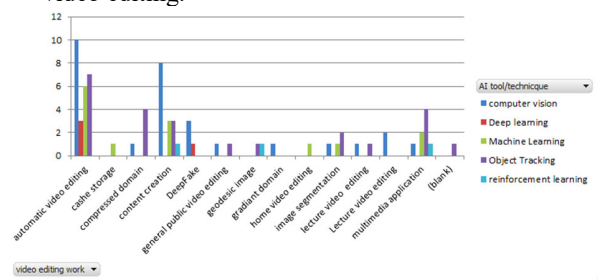

Fig. 2. Use of AI techniques in video editing

Above mentioned graphs help us to determine which AI technique is used in the particular research area of video editing. As mentioned in table 3 video editing categories have different trends of using AI tools and technologies. The trend of using AI techniques is higher in automatic video editing than in other research areas of video editing. Furthermore, automatic video editing use most branches of AI, in their research. The use of computer vision has been widely adopted by researchers of automatic video editing. Afterward, object tracking and then machine learning also has been widely used in automatic video editing. Deep learning is an emerging technology not widely adopted for automatic video editing research. I look into other video editing research areas; we can come to know that the compression domain is widely using computer vision as an AI tool. Multimedia video editing is a research area of video editing that also uses a variety of AI techniques such as computer vision, machine learning, object tracking, reinforcement learning and deep learning.

- RQ3: the purpose of this question is to evaluate the worth of research publications on video editing. We choose only four types of venues that are conference, journal, symposium, and workshop. Since from start our research focuses on well-known data sources, we assume journals of that data source most worthy publication. Afterward, the symposium then the conferences, and at the bottom, we assume workshop papers have lowest work.
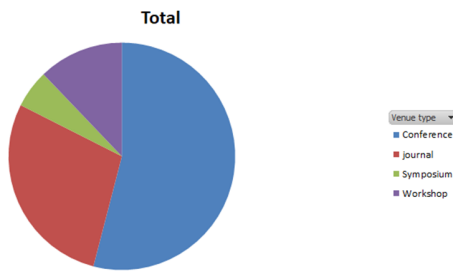
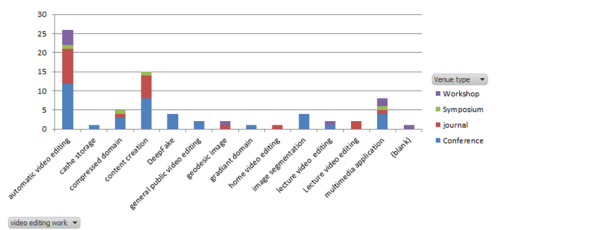Fig 3. Video editing publications venue wise

Fig 4. Video editing research publication according to their venue

From this graph, we can come to know that automatic video editing is the domain of video editing that uses almost all kinds of venues, but conferences are most prominent in them. While, if we discuss content creation, the majority of paper in this research area has been published in Journals that are likely to be most worthy. Deepfake, a research area of video editing, is not much popular, yet has the significance that all publications of this area are published in well-known journals. Here we did not take into account the impact factor of journals because all journals are ISI indexed and no stand-alone publication come in our search.

- RQ4: this is final question and will be more fruitful for new researcher in video editing domain. This question tries to find out research gap in our particular research area. As shown in Fig 5 we can see a growing trend of AI in video editing. This graph shows in which frequency, video editing research areas using AI tool.
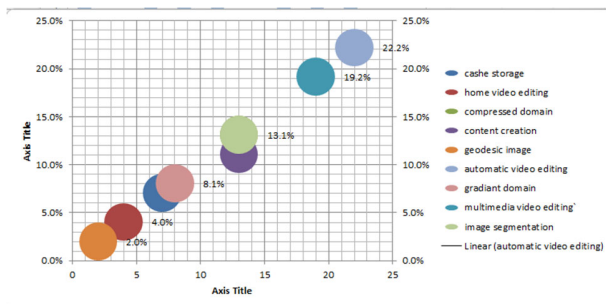

Fig 5. Bubble char of video editing categories research areas

- From this graph, we can see that automatic video editing is an area of video editing that uses AI tools and technology the most. Afterward, multimedia video editing is using AI tools. Hence new researchers can find gaps in using AI tools and technology in video editing research domain.

4. Conclusion

We present novel research in the video editing domain that tried to find uses of AI in video editing. AI is the most powerful tool of today's research world. It is not only applicable in the specific computing world rather it is widely used in all aspects of life. The purpose of using AI is to produce intelligent and efficient work. Growing AI technology like machine learning, computer vision, deep learning, etc are playing their role in the advancement of technology for better use of technology. In this research, we propose some research questions that are purely related to using of AI in video editing. We extracted data from well-reputed digital libraries and performed a statically approach to find answers to our research question. Our research will be very helpful for future researchers and industrialists of video editing to find research gaps and hence will be able to find specific research areas and can start their research. Our research will also be a big milestone for new doctorate students to determine their particular research area in video editing. From our research, we find that automatic video editing and multimedia video editing are the fields that are increasingly using AI tools. Furthermore, we also conclude that computer vision and machine learning are the tools of AI that are mostly adopted in video editing research. Not only above mentioned areas but also there are a lot more areas, that is mentioned in our paper in which new researcher can start their research. In the future, we have aimed to narrow down our research and more specifically find the use of AI in one of the above-mentioned video editing subdomains.

References

[1] Dancyger, K.: *Q site for the technique of film and video editing: History, theory, and practice*. Focal Press. (2014).

[2] Outtagarts, A., & Mbodj, A.: A Cloud-Based Collaborative and Automatic Video Editor. In *2012 IEEE International Symposium on Multimedia (ISM)*. IEEE. https://doi.org/10.1109/ism.2012.78 (2012).

[3] Soe, T. H. (2021). AI video editing tools. What editors want and how far is AI from delivering? In *arXiv [cs.HC]*. http://arxiv.org/abs/2109.07809

[4] Leake, M., Davis, A., Truong, A., & Agrawala, M. (2017). *Computational video editing for dialogue-driven scenes*. ACM Transactions on Graphics, vol. *36*(4), pp.1–14. https://doi.org/10.1145/3072959.3073653

[5] Pavel, A., Reed, C., Hartmann, B., & Agrawala, M.: *Video digests: A browsable, skimmable format for informational lecture videos*. Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology - UIST '14. doi:10.1145/2642918.2647400 (2014).

[6] Liu, T., & Kender, J. R.: *Lecture videos for E-learning: Current research and challenges*. IEEE Sixth International Symposium on Multimedia Software Engineering. https://doi.org/10.1109/mmse.2004.48(2005).

[7] Wu, Y., Mei, T., Xu, Y.-Q., Yu, N., & Li, S.: MoVieUp: *Automatic Mobile Video Mashup*. IEEE Transactions on Circuits and Systems for Video Technology: A Publication of the Circuits and Systems Society, vol. 25(12), pp. 1941–1954. https://doi.org/10.1109/tcsvt.2015.2416554(2015).

[8] Frey, N., Chi, P., Yang, W., & Essa, I.: *Automatic Non-Linear Video Editing Transfer*. arXiv e-prints, arXiv-2105. (2021).

[9] Rowe, R. S.: *Remote non-linear video editing*. SMPTE journal, vol.109(1), pp.23-25. (2000).

[10] Sumec, S. (2006). Multi camera automatic video editing. In *Computational Imaging and Vision* (pp. 935–945). Kluwer Academic Publishers.

[11] Hada, Y., Ogata, H., & Yano, Y.: *Video-based language learning environment using an online video-editing system*. Computer Assisted Language Learning, vol. 15(4), pp. 387–408. https://doi.org/10.1076/call.15.4.387.8273 (2002).

[12] Outtagarts, A., Squedin, S., & Martinot, O.: Cloud-based automatic video editing using keywords. In *E-Business and Telecommunications* (pp. 228–241). Springer Berlin Heidelberg. (2014).

[13] Wang, M., Lyu, X.-Q., Li, Y.-J., & Zhang, F.-L.: *VR content creation and exploration with deep learning: A survey*. Computational Visual Media, vol. 6(1), pp. 3–28. https://doi.org/10.1007/s41095-020-0162-z (2020).

[14] Feng, J., Li, S., Li, X., Wu, F., Tian, Q., Yang, M.-H., & Ling, H.: *TapLab: A Fast Framework for Semantic Video Segmentation Tapping into Compressed-Domain Knowledge*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 1. https://doi.org/10.1109/tpami.2020.3024646 (2020).

[15] Kwok, A. O. J., & Koh, S. G. M.: D*eepfake: a social construction of technology perspective*. Current Issues in Tourism, vol. 24(13), pp.1798–1802. https://doi.org/10.1080/13683500.2020.1738357 (2021).

[16] Girgensohn, A., Boreczky, J., Chiu, P., Doherty, J., Foote, J., Golovchinsky, G., ... & Wilcox, L.: *A semi-automatic approach to home video editing.* In Proceedings of the 13th annual ACM symposium on User interface software and technology, pp. 81-89. (2000, November).

[17] Criminisi, A., Sharp, T., Rother, C., & P'erez, P.: *Geodesic image and video editing.* ACM Transactions on Graphics, vol. 29(5), pp. 1–15. https://doi.org/10.1145/1857907.1857910 (2010).

[18] Wang, H., Xu, N., Raskar, R., & Ahuja, N.: *Videoshop: A new framework for spatio-temporal video editing in gradient domain*. Graphical Models, vol. 69(1), pp. 57–70. https://doi.org/10.1016/j.gmod.2006.06.002(2007).

[19] Pitas, I.: *Digital image processing algorithms and applications*. John Wiley & Sons. (2000).

Igor Bieda, Department of Theory and Technology of Programming
Faculty of Computer Science and Cybernetics
Taras Shevchenko National University of Kyiv
Akademika Glushkova Ave., 4-D, Kyiv, Ukraine, 03680
tp@infosoft.ua

Taras Panchenko, Department of Theory and Technology of Programming
Faculty of Computer Science and Cybernetics
Taras Shevchenko National University of Kyiv
Akademika Glushkova Ave., 4-D, Kyiv, Ukraine, 03680
tp@infosoft.ua