

MINISTERUL EDUCAȚIEI ȘI CERCETĂRII AL REPUBLICII MOLDOVA

Universitatea Tehnică a Moldovei

Facultatea Calculatoare, Informatică și Microelectronică

Departamentul Ingineria Software și Automatică

**Admis la susținere
Șef departament:
FIODOROV Ion dr., conf.univ.**

„___” _____ 2024

DEZVOLTAREA UNUI EDITOR VIDEO CU FUNCȚIONALITĂȚI DE INTELIGENȚĂ ARTIFICIALĂ

Teză de master

Student: _____ **Rusu Dionisie, TI-221M**

Coordonator: _____ **Bolun Ion, prof. univ.**

Consultant: _____ **Cojocaru Svetlana, asist.univ.**

Chișinău, 2024

REZUMAT

la teza de master „Dezvoltarea unui editor video cu funcționalități de inteligență artificială” a masterandului din grupa TI-221M, programul de studii „Tehnologia informației”,
RUSU Dionisie

Teza de cercetare de master propusă explorează aplicarea inteligenței artificiale în domeniul editării video. Lucrarea cuprinde introducerea, șase capitole, concluzii, bibliografia din 18 titluri, 12 anexe și este perfectată pe 87 pagini de text de bază.

În Capitolul 1 sunt descrise aspecte introductive și contextul tematic. Se prezintă cadrul tematic general, starea actuală și perspectivele viitoare în domeniul inteligenței artificiale aplicate editării video. De asemenea, se evidențiază tehnologiile AI utilizate în manipularea datelor video și audio, iar problema de cercetare este definită pentru a ghida întreaga investigație.

Capitolul 2 se axează pe o cercetare teoretică detaliată a clasificării conținutului video și audio, focalizându-se asupra seturilor de date utilizate pentru modele. Se analizează extensiv funcționalitățile editării video automate non-AI, cum ar fi tăierea și asamblarea clipurilor, adăugarea de efecte speciale și slow-motion. De asemenea, se explorează editarea automată, segmentarea automată a scenelor și posibilitatea adaptării conținutului video la conținutul audio, evidențiind tehnologii existente în aceste domenii.

Capitolul 3 explorează detaliat prototipul editorului VIDEO-AI, evidențiind funcționalitățile esențiale ale acestuia. Se enumeră și analizează instrumentele necesare pentru construcția prototipului, iar metodele de formare și setul de date pentru antrenarea modelelor. În încheiere, se discută despre potențialele aplicații ale editorului în domeniul educației și în rețelele sociale, oferind o perspectivă asupra impactului său în diverse contexte.

Proiectarea arhitecturii editorului VIDEO-AI este detaliată în Capitolul 4, acoperind aspecte comportamentale și structurale. Se descrie în detaliu editorul VIDEO-AI, modelarea fluxurilor de lucru, stările și tranzițiile în funcționarea sa, precum și interacțiunile și fluxurile de date. Arhitectura componentelor software și interdependențele dintre module sunt, de asemenea, discutate.

Capitolul 5 aduce în prim-plan analiza algoritmică pentru extragerea semantică VIDEO-AI. Aici se explorează și analizează algoritmi de segmentare video, evaluând și comparând algoritmi de analiză a conținutului video.

Implementarea tehnologică a prototipului editorului VIDEO-AI este prezentată în Capitolul 6. Acesta acoperă integrarea tehnologiilor de înțelegere a conținutului video, inclusiv clasificarea conținutului vizual și audio. Se detaliază realizarea automată a editării conținutului video, automatizarea montajului video prin analiza audio și implementarea unei interfețe adaptative pentru editarea video.

Prin abordarea detaliată a acestor aspecte, teza oferă o privire cuprinzătoare asupra aplicării inteligenței artificiale în domeniul editării, evidențiind contribuțiile și inovațiile aduse de VIDEO-AI.

ABSTRACT

of the master's thesis „**Development of a video editor with artificial intelligence features**” by the master's student from the TI-221M group, „Information Technology” study program,

RUSU Dionisie

The proposed master's research thesis explores the application of artificial intelligence in the field of video editing. The work includes an introduction, six chapters, conclusions, a bibliography with 18 titles, 12 annexes, and consisting of 87 pages of core text.

Chapter 1 describes introductory aspects and the thematic context. It presents the general thematic framework, the current state, and future prospects in the field of artificial intelligence applied to video editing. It also highlights the AI technologies used in video and audio data manipulation, defining the research problem to guide the entire investigation.

Chapter 2 focuses on a detailed theoretical research of video and audio content classification, with a focus on the datasets used for models. It extensively analyzes non-AI automated video editing functionalities, such as cutting and assembling clips, adding special effects, and slow motion. It also explores automatic editing, automatic scene segmentation, and the possibility of adapting video content to audio content, highlighting existing technologies in these areas.

Chapter 3 explores in detail the prototype of the VIDEO-AI editor, highlighting its essential functionalities. It lists and analyzes the tools required for building the prototype, along with the training methods and dataset used for model training. In conclusion, it discusses the potential applications of the editor in education and social networks, offering a perspective on its impact in various contexts.

The design of the VIDEO-AI editor's architecture is detailed in Chapter 4, covering behavioral and structural aspects. The VIDEO-AI editor's workflow modeling, states, transitions in its operation, as well as interactions and data flows, are described in detail. The architecture of software components and the interdependencies between modules are also discussed.

Chapter 5 brings to the forefront the algorithmic analysis for VIDEO-AI semantic extraction. In this chapter, video segmentation algorithms are explored and analyzed, including the evaluation and comparison of video content analysis algorithms.

The technological implementation of the VIDEO-AI editor prototype is presented in Chapter 6. It covers the integration of video content understanding technologies, including visual and audio content classification. The automatic implementation of video content editing, video montage automation through audio analysis, and the implementation of an adaptive interface for video editing are detailed.

By addressing these aspects in detail, the thesis provides a comprehensive view of the application of artificial intelligence in the field of editing, highlighting the contributions and innovations brought about by VIDEO-AI.

CUPRINS

INTRODUCERE	9
1 INTELIGENȚA ARTIFICIALĂ ÎN EDITAREA VIDEO – GENERALITĂȚI	10
1.1 Cadrul tematic, stare și perspective în domeniu.....	11
1.2 Tehnologii AI în operarea datelor video și audio.....	13
1.3 Definirea problemei de cercetare	15
2 EDITAREA VIDEO: ASPECTE TEORETICE	17
2.1 Clasificarea conținutului video și audio	18
2.2 Funcționalitățile editării video automate non-AI	21
2.3 Metode de editare automată a secvențelor video	23
2.4 Segmentarea automată a scenelor video.....	25
2.5 Adaptarea secvențelor video pe baza analizei conținutului audio.....	28
3 PROTOTIPUL EDITORULUI VIDEO-AI: ASPECTE CONCEPTUALE	31
3.1 Funcționalități ale editorului VIDEO-AI	32
3.2 Instrumente pentru realizarea editorului VIDEO-AI	33
3.3 Setul de date pentru antrenarea modelelor editorului VIDEO-AI.....	35
3.4 Posibilități de aplicare a editorului VIDEO-AI.....	36
4 PROIECTAREA ARHITECTURII EDITORULUI VIDEO-AI	38
4.1 Proiectarea comportamentală	39
4.1.1 Descrierea generală a editorului VIDEO-AI	40
4.1.2 Modelarea fluxurilor de lucru	42
4.1.3 Stările și tranzițiile în funcționarea editorului VIDEO-AI	43
4.1.4 Interacțiunile și fluxurile de date	45
4.2 Proiectarea structurală a componentelor	46
4.2.1 Arhitectura componentelor software	47
4.2.2 Interdependențele dintre module	49
4.2.3 Modelarea implementării tehnologice	50

5 ANALIZA ALGORITMICĂ PENTRU EXTRAGEREA SEMANTICĂ VIDEO-AI	51
5.1 Explorarea și analiza algoritmilor de segmentare video	52
5.2 Evaluarea și compararea algoritmilor de analiză a conținutului video	54
6 IMPLEMENTAREA TEHNOLOGICĂ A PROTOTIPULUI EDITORULUI VIDEO-AI.....	56
6.1 Integrarea tehnologiilor de înțelegere a conținutului video	57
6.1.1 Clasificarea conținutului vizual	58
6.1.2 Clasificarea conținutului audio	64
6.1.3 Modele de înțelegere a conținutului video	66
6.2 Realizarea automată a editării conținutului video	71
6.3 Automatizarea montajului video prin analiza audio	76
6.4 Interfață adaptivă pentru editarea video	84
CONCLUZII.....	86
BIBLIOGRAFIE.....	88
ANEXA A Imaginea generală asupra sistemului.....	90
ANEXA B Stările de tranzacție ale sistemului	92
ANEXA C Descrierea scenariilor de utilizare a editorului	95
ANEXA D Fluxurile de mesaje și legăturile dintre componentele sistemului.....	98
ANEXA E Descrierea structurii statice a sistemului	99
ANEXA F Relațiile de dependență între componentele sistemului	100
ANEXA G Modelarea echipamentelor mediului de implementare.....	102
ANEXA H Compararea algoritmilor de captioning.....	103
ANEXA I Rezultatul procesului de înțelegere a conținutului video.....	104
ANEXA Î Procesul de antrenare a unui model în cadrul la Techable Machine	105
ANEXA J Modulul de procesare a hărții de editare	106
ANEXA K Modulul de asamblare și salvare a unui edit în baza muzicii	107

INTRODUCERE

În era digitală în continuă evoluție în care trăim astăzi, conținutul video a devenit o componentă esențială a comunicării noastre cotidiene. De la platforme de socializare și site-uri de partajare a conținutului, până la producția de conținut profesional în domenii precum cinematografia și publicitatea, videoclipurile au dobândit o importanță fără precedent. Cu toate acestea, crearea și editarea conținutului video rămân adesea procese complexe și consumatoare de timp, necesitând cunoștințe tehnice avansate și instrumente sofisticate.

În acest context, Inteligența Artificială (AI) a apărut ca un element revoluționar, schimbând fundamental modul în care abordăm producția și editarea conținutului video. Prin intermediul AI, avem acum posibilitatea de a crea și îmbunătăți conținutul video în moduri inimaginabile anterior, automatizând procesele și aducând inovații semnificative în domenii precum post-producția, publicitatea digitală și educația.

Această cercetare se concentrează în special pe intersecția dintre două domenii-cheie: inteligența artificială și producția video. Inteligența artificială reprezintă capacitatea mașinilor de a învăța și de a efectua sarcini specifice fără intervenția umană. Această abordare a revoluționat numeroase industrii, iar industria video nu face excepție. Editorii video cu funcționalități AI sunt dezvoltați pentru a automatiza și îmbunătăți procesele de editare video, oferind astfel creatorilor de conținut posibilitatea de a produce materiale de înaltă calitate cu mai puțin efort și timp.

Examinând modul în care inteligența artificială poate fi integrată în dezvoltarea unui editor video și evaluând impactul acestei integrări asupra procesului de editare video, precum și asupra calității produselor finisate, se poate aborda o serie de aspecte cheie, inclusiv tehnologiile AI utilizate în producție și editare video, avantajele și dezavantajele utilizării acestor tehnologii și modul în care acestea influențează industria video în ansamblu.

Un alt aspect important al acestui studiu este explorarea aplicațiilor practice ale unui editor video cu funcționalități AI în diverse domenii, precum industria filmului, rețelele de socializare și educația. Prin implementarea acestei tehnologii în aceste domenii, cercetarea vizează să dezvăluie modul în care aceste tehnologii pot schimba și îmbunătăți procesele existente, aducând inovații semnificative în modurile în care producem și consumăm conținut video.

Inteligența artificială și producția video sunt două forțe motrice care converg pentru a redefine modul în care comunicăm și interacționăm cu lumea video. Cu evoluția continuă a acestor tehnologii, sunt așteptate schimbări semnificative în modul în care are loc producerea, consumul și conectarea între oameni prin intermediul conținutului video. Astfel, inteligența artificială continuă să fie un motor de inovație și transformare în era digitală, deschizând noi orizonturi pentru creativitatea și eficiența umană.

BIBLIOGRAFIE

- [1] A. Reena, „How AI in Audio and Video Editing is Saving Time and Effort”, Podcastle Blog. Data accesării: 16 septembrie 2023. [Online]. Disponibil la: <https://podcastle.ai/blog/benefits-of-ai-in-audio-and-video-editing/>
- [2] Z. Xinrong, Y. Li, Y. Han, și J. Wen, „AI Video Editing a Survey - Review AI Video Editing: a Survey Xinrong Zhang 1 , Yanghao Li2,†, - Studocu”. Data accesării: 23 septembrie 2023. [Online]. Disponibil la: <https://www.studocu.com/ph/document/dmc-college-foundation/science-technology-and-society/ai-video-editing-a-survey/41148700>
- [3] „Edit videos faster with artificial intelligence in Adobe Premiere Pro”. Data accesării: 18 septembrie 2023. [Online]. Disponibil la: <https://www.adobe.com/products/premiere/ai-video-editing.html>
- [4] R. Kundu, „Video Classification: Methods, Use Cases, Tutorial”. Data accesării: 18 septembrie 2023. [Online]. Disponibil la: <https://www.v7labs.com/blog/video-classification-guide>
- [5] „UCF101 - Action Recognition Data Set”. Data accesării: 21 septembrie 2023. [Online]. Disponibil la: <https://www.crcv.ucf.edu/data/UCF101.php>
- [6] „Hueihan Jhuang. HMDB: a large human motion database”. Data accesării: 21 septembrie 2023. [Online]. Disponibil la: <https://serre-lab.clps.brown.edu/resource/hmdb-a-large-human-motion-database/>
- [7] „YouTube-8M: A Large and Diverse Labeled Video Dataset for Video Understanding Research”. Data accesării: 21 septembrie 2023. [Online]. Disponibil la: <https://research.google.com/youtube8m/>
- [8] J. Fumo, „Types of Machine Learning Algorithms You Should Know”, Medium. Data accesării: 24 septembrie 2023. [Online]. Disponibil la: <https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861>
- [9] A. Roy, „A Guide to Image Captioning”, Medium. Data accesării: 27 septembrie 2023. [Online]. Disponibil la: <https://towardsdatascience.com/a-guide-to-image-captioning-e9fd5517f350>
- [10] Vangie Beal, „Video Editing”. Data accesării: 23 septembrie 2023. [Online]. Disponibil la: <https://www.webopedia.com/definitions/video-editing/>
- [11] Khushboo, „Video Scene Detection and Classification: PySceneDetect, Places365 and Mozilla DeepSpeech Engine”, Medium. Data accesării: 24 septembrie 2023. [Online]. Disponibil la: <https://zwolf12.medium.com/video-scene-detection-and-classification-pyscenedetect-places365-and-mozilla-deepspeech-engine-51338e3dbacc>
- [12] C. Shixing, N. Xiaohan, F. David, Z. Dongqing, și B. Vimal, „Automatically identifying scene boundaries in movies and TV shows”, Amazon Science. Data accesării: 24 septembrie 2023. [Online]. Disponibil la: <https://www.amazon.science/blog/automatically-identifying-scene-boundaries-in-movies-and-tv-shows>

- [13] „PySceneDetect / Инструмент для автоматического разделения видео на отдельные клипы”. Data accesării: 24 septembrie 2023. [Online]. Disponibil la: [//zenway.ru/page/pyscenedetect](http://zenway.ru/page/pyscenedetect)
- [14] „Audio Analysis With Machine Learning: Building AI-Fueled So”, AltexSoft. Data accesării: 27 septembrie 2023. [Online]. Disponibil la: <https://www.altexsoft.com/blog/audio-analysis/>
- [15] „Get image descriptions using visual captioning | Vertex AI”, Google Cloud. Data accesării: 16 noiembrie 2023. [Online]. Disponibil la: <https://cloud.google.com/vertex-ai/docs/generative-ai/image/image-captioning>
- [16] R. Alec, S. Ilya, W. Jong, și K. Gretchen, „CLIP: Connecting text and images”. Data accesării: 11 noiembrie 2023. [Online]. Disponibil la: <https://openai.com/research/clip>
- [17] „Object Detection | TensorFlow Hub”, TensorFlow. Data accesării: 21 noiembrie 2023. [Online]. Disponibil la: https://www.tensorflow.org/hub/tutorials/object_detection
- [18] „Audio classification guide for Python”, Google for Dev. Data accesării: 8 noiembrie 2023. [Online]. Disponibil la: https://developers.google.com/mediapipe/solutions/audio/audio_classifier/python