AI4media

ARTIFICIAL INTELLIGENCE FOR THE MEDIA AND SOCIETY

ROADMAP ON AI TECHNOLOGIES & APPLICATIONS FOR THE MEDIA INDUSTRY

SECTION: "AI FOR NEWS PRODUCTION"

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

info@ai4media.eu www.ai4media.eu

Authors	**Sabino Metta** (Rai – Radiotelevisione Italiana) **Alberto Messina** (Rai – Radiotelevisione Italiana)

This report is part of the deliverable D2.3 - "*AI technologies and applications in media: State of Play, Foresight, and Research Directions*" of the AI4Media project.

You can site this report as follows:

F. Tsalakanidou et al., Deliverable 2.3 - AI technologies and applications in media: State of play, foresight, and research directions, AI4Media Project (Grant Agreement No 951911), 4 March 2022

This report was supported by European Union´s Horizon 2020 research and innovation programme under grant number 951911 - AI4Media (A European Excellence Centre for Media, Society and Democracy).

The information and views set out in this report are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf.

AI for news production

Current status

News is more than just a tangible product. It is a public good that contributes to the pillars of democracy. The availability of news allows citizens to be informed about issues that concern them so that they can make better decisions. So, it is fundamental that news is produced professionally and ethically. On the other hand, the digitalisation of media is supposed to increase the tension between the production of news as a public good vs. its delivery as a commodity. Nowadays, news is made readily available and accessible on multiple sources: online, radio, television, social networks, etc. Information is continuously produced by large groups of professional journalists and communication experts. It is also ourselves and our own network of people that continuously produce and propagate information simply by sharing them. In this media ecosystem the "time" variable plays a crucial role. Journalists are focusing on producing news quicker with respect to reporting more in-depth news and the entire news production chain is stressed to work at high speed.

Figure 1: A Rai News 24 reporter reporting from London.

Technology in the newsroom has helped to "speed up" news production and distribution, including personalisation, marketing, finding audiences, understanding user behaviour, monetisation/subscriptions[1]. Technological advancements are aiding journalists to deal with the pressure in the newsrooms. By using light devices (mobiles, tablets, etc.) journalists can record and report on news from wherever they are and easily upload items onto news sites as well as social media. Tools allow journalists to search and retrieve specific contents stored in newsroom archives thus enriching their stories. The speed of the internet enables journalists to find and get data without having to leave the newsroom. Furthermore, the internet allows journalists to gather and analyze possible comments provided from end users thus inspiring new investigations. In addition, in the age of digital and social media, visual communication has sky-rocketed. Images are powerful, we know. Visuals quickly transmit information to our brain and trigger our emotions thus stimulating our engagement. This can be particularly difficult for TV

[1] C. Beckett. "New powers, new responsibilities. A global survey of journalism and artificial intelligence". November 18th 2019, JournalismAI | Research, London School of Economics and Political Science.

news production where the storytelling is based on the availability of images and videos. When a chronicle fact occurs, live footage allows to tell and to show what has been happening. In case of unexpected events (like natural disasters: floods, earthquakes, pandemic, etc.), newsroom's correspondents need to readily publish and update fresh news in a very short timeframe (Figure 1).

To manage very low bandwidth video transmission, ***AI technologies can increase the efficiency of video compression***. Algorithms calculate bitrate to optimise bandwidth usage while maintaining an appropriate level of quality.

Audiovisual archives can enrich and empower storytelling. When a flood (or any other natural disaster) occurs, audiovisual archives can be used for documenting environmental and urban conditions of a location before the event occurred. In this way, damages caused by the flood can be highlighted by images before and after the disaster has occurred. Currently, media organisations face several issues regarding access to their archives: different and time-changing A/V formats, video content rights, different definitions (standard, FullHD, 4K, etc.), several types of compression (MPEG2, MPEG4, AVC, etc.), metadata, wrappers (MXF, Quicktime), large A/V storages (petabytes). Content Management Systems oversee making video content findable, referenceable and reusable.

On the other hand, ***AI technologies can enhance video archives***. These technologies can efficiently automate different tasks: metadata enrichment, splitting of long videos, object and scene identification, facial recognition, audio transcription, on-screen text extraction, etc. Machine learning technologies can automate the generation of searchable metadata tags by timecode (about actions, places and things identifiable within the content, etc.) thus enhancing search and discovery tasks in content libraries. In addition to the opportunity to properly document occurring events, the extraction of generic scenes within video segments (e.g. a rainy day, a landslide in the mountains, etc.) can be opportunistically exploited for evoking ad-hoc context and enhancing storytelling[2,3].

Recent developments in deep learning models are revolutionising the traditional workflow for video creation. In 2018, China's state news agency Xinhua developed AI anchors through machine learning technologies. Voice, facial movements, and gestures of real-life broadcasters have been simulated[4]. Currently, there exist several AI video generator tools that provide virtual human presenters able to convey flawless human-like voice speeches for different use cases (digital marketing, corporate communications, employee training, etc.). On the other hand, "deep fake" videos can fool even expert personnel. Fortunately, plenty of software tools are available for digital image forensics. These tools support journalists to detect if a photo or a video has been manipulated or faked. This is particularly important in case of user-generated

[2] Kenneth E. Foote, "To Remember and Forget: Archives, Memory, and Culture", The American Archivist, Vol. 53, No. 3 (Summer, 1990), pp. 378-392 (15 pages), Published By: Society of American Archivists

[3] Reuters Staff, "Reuters applies AI technology to 100 years of archive video to enable faster discovery, supported by Google DNI", AUGUST 13, 2020, Reuters web page: https://www.reuters.com/article/rpb-lavita-video-archive-idUSKCN2591VO

[4] Lily Kuo, "World's first AI news anchor unveiled in China", 9th November 2018, The Guardian, https://www.theguardian.com/world/2018/nov/09/worlds-first-ai-news-anchor-unveiled-in-china

content (UGC). Furthermore, AI-based technologies can solve several problems: denoising, super-resolution, etc.[5].

Research challenges

News production is a quite wide process involving several tasks and issues. By focusing on the production of audiovisual news (e.g. TV news programs), research challenges have still to be addressed in the following areas:

Video compression. Video standards have existed since the early 1990s. The demand for multimedia has increased and a huge amount of data is being produced. So, efficient communications need to advance compression technology. Neural network "encoders" are challenging the conventional approach. AI-based compression can provide the same level of visual quality with fewer bits[6,7,8,9,10,11].

Information retrieval. It is a hard task for computer systems to analyze, understand and represent contents in natural language. Neural networks have been applied to this task and need more investigation and trial in the professional media domain. AI and ML technologies (in particular "deep learning" technologies) can support different tasks:

- ***Natural language processing***. This allows to better understand and model the user information need by exploiting the huge amount of corpora of information.
- ***Image/video recognition***. This allows to extract features and search a multimedia corpus of information (e.g. objects and entities involved rather than just pixel and colour-related information).
- ***Knowledge representation***. This allows to build better data structures and search algorithms to identify meaning, synonyms, and relations between terms and concepts.
- ***Learning***. This allows to learn relevance ranking functions, to classify query intent and documents and to offer personalised results.

Automatic Speech Recognition (ASR), also known as Speech to Text. AI-powered speech recognition technology has the power to convert speech content to text and to recognise an individual based on their voice command.

[5] U. Schmidt and R. Stefan. "Shrinkage Fields for Effective Image Restoration". IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014.
[6] D. Ding, Z. Ma, Di Chen, Q. Chen, Z. Liu, and F. Zhu. "Advances in Video Compression System Using Deep Neural Network: A Review and Case Studies". January 2021, arXiv:2101.06341v1
[7] E. Çetinkaya, H. Amirpour, C. Timmerer, and M. Ghanbari, "FaME-ML: Fast Multirate Encoding for HTTP Adaptive Streaming Using Machine Learning", IEEE International Conference on Visual Communications and Image Processing (VCIP), 2020.
[8] Cisco Annual Internet Report, 2018–2023
[9] M. Nagel, M. van Baalen, T. Blankevoort, M. Welling. "Data-Free Quantization through Weight Equalisation and Bias Correction." IEEE International Conference on Computer Vision (ICCV), 2019.
[10] A. Kuzmin, M. Nagel, S. Pitre, S. Pendyam, T.Blankevoort, M. Welling. "Taxonomy and Evaluation of Structured Compression of Convolutional Neural Networks.", 2019, CoRR, arXiv:1912.09802
[11] QUALCOMM, "How AI research is enabling next-gen codecs" (2021): https://www.qualcomm.com/news/onq/2021/07/14/how-ai-research-enabling-next-gen-codecs

Generative Adversarial Networks (GANs). Deep learning technologies can synthesise highly convincing images and voices. However, natural video generation is still lagging.

Societal and media industry drivers

Vignette: AI-enabled live news coverage of a flood event in Sarno, Italy

Sarno (Italy). It's been raining for around 72 hours. Maria Luisa works as a journalist for RAI, the Italian national broadcaster. Today, due to some inconveniences along the road, she decided to work from home. At 3:14 pm (local time), while she is in front of her PC, a roar suddenly startles her. Looking out of the window she is completely shocked. It’s raining heavily now and the roar's cause is evident: a landslide has broken off the hills right in front of her house. Immediately, Maria Luisa warns the police about the occurred event and a few seconds later she alerts her newsroom. To save time, Maria Luisa uses her mobile phone for video calling her colleagues and shows the landslide. Likely due to the bad weather conditions, the performance of Internet connection is low. Nevertheless, AI-powered algorithms optimise bandwidth usage while maintaining an appropriate level of quality. So, Maria Luisa and colleagues of her, decide to go live on the "Rai News 24" channel. From the balcony of her house, Maria Luisa can broadcast “live” the severity of the event: it rains a lot and a landslide just broken off and is visible on the hills.

Within a few minutes, the RAI control room (central newsroom) is overwhelmed by a multitude of reports and videos shot from local citizens thus witnessing what is happening at Sarno and at surrounding areas. Supported by AI tools, RAI's newsroom can manage this huge amount of information and automate different tasks: metadata extraction, video enhancing, scene identification, etc. In this way, RAI can broadcast and update in real time a map of Sarno and surroundings showing the ongoing situation (streets and/or bridges closed, dangerous areas, etc.). Videos sent from citizens can be verified, selected and broadcasted live. So, citizens receive live fresh information about what is going on in their city.

Due to bad weather conditions, no drones or any other aircrafts can fly and take aerial images. Nevertheless, available Digital Surface Models (DSMs) of that area allow technicians within the newsroom to build a 3D visualisation. AI-powered tools allow newsroom staff to generate a virtual simulation of the landslide that occurred a few minutes ago. This simulation is live broadcasted thus enhancing the storytelling. This virtual representation highlights the actual risks associated to possible further landslides that might break off in the next hours. Live news broadcast contributes to speed up the reaction of citizens. Local authorities can easily begin to evacuate several areas. At 6:04 pm (local time), a strong roar occurs: a second big landslide breaks off from the Pizzo D'Alvano mountain nearby. Two million cubic metres of mud fall upon Sarno. Fortunately, most houses had been evacuated and many human lives have been saved.

The day after it is not raining anymore. By searching in internal archives, RAI journalists do not find any similar event to have occurred in that area. The only aerial footage of Sarno is that recorded from a helicopter for the Giro d'Italia (cycling race) some years before. These footages document the urban and environmental conditions before the disaster. AI technologies can analyse these videos and extract relevant information (e.g., flight path, altitude, speed, camera angle, etc.). This information can be exploited for managing flight and shooting parameters of a drone that can document the current situation. By reproducing compatible manoeuvres of the helicopter, the drone's footage can highlight the damages caused from the event that just occurred. The day after, Maria Luisa shows the area covered by the mud. She can investigate the reasons and the responsibilities why the event has occurred.

Future trends for the media sector

Despite the concerns regarding the harmful use of these technologies (e.g. "deep fakes"), AI-powered tools will enhance the production of news. Developed for solving repetitive tasks, these technologies will be more and more used for freeing up time for novel "creative" skills.

By enhancing video compression tasks, AI technologies will enhance the streaming of videos on mobile devices. Real-time and high-quality video news will be more and more watched on small and portable devices. These growing consumer expectations will be met by the future availability of 5G technologies that will increase the speed of the wireless networks.

Newsrooms will elaborate on more and more amounts of information, related to facts that occurred locally and worldwide. Journalists will rely on AI-powered tools for news reporting and/or investigative projects[12]. AI technologies will be used for automatic content indexing and retrieval, detection (e.g., people, places, objects, concept, etc.), text/video summaries generation, video editing, generation of synthesised video (e.g., virtual presenters), translation in different languages. For a media company, only contents really matter. So, news will be delivered on any existing platform. Even on a still not fully clear one, like the Metaverse.

Goals for next 10 or 20 years

AI technologies have shown to evolve so rapidly that it is difficult to predict what will happen over the next 10-20 years. Newsrooms will rely on AI technologies for enhancing and changing internal workflows. Most repetitive tasks (video editing, summary generation, translation and delivery in different languages, etc.) will be performed by software tools. This opportunity will allow journalists to focus mainly on ongoing investigation/projects. To better meet consumer expectations, newsrooms will establish new synergies with citizens. AI-powered technologies will support information (and video) exchange guaranteeing ethical and professional aspects. To enhance internal training, newsrooms will foster new partnerships and joint projects with scientists and researchers. By the usage of advanced technologies, newsroom personnel will perform more investigations thus returning a deeper information to final users.

[12] S. McGoey, "A decade of digital evolution to help reporting revolutions at ICIJ", May 25th 2021, ICIJ website, https://www.icij.org/investigations/panama-papers/a-decade-of-digital-evolution-to-help-reporting-revolutions-at-icij/

AI4media

ARTIFICIAL INTELLIGENCE FOR THE MEDIA AND SOCIETY

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

info@ai4media.eu www.ai4media.eu