Exploring AI-Enabled Use Cases for Societal Security and Safety

Hoang Long Nguyen, Minsung Hong, Rajendra Akerkar[*]

Big Data Research Group, Western Norway Research Institute
P.O.Box 163, NO-6851 Sogndal, Norway
{hln, msh, rak}@vestforsk.no

Abstract

Global momentum around Artificial Intelligence (AI) for social good is growing. AI opens a new perspective to maintain public security and safety by providing investigative assistance with a human-grade precision. Quantitative methods might always not be a correct evaluation of the AI techniques due to the characteristics of the societal security domain. Therefore, we also need qualitative research methods in relevant use cases. The paper presents two AI-enabled use cases on information validation and surveillance enhancement with the support of AI algorithms.

Introduction

Researches in the field of societal security and safety are related to critical events that can cause a threat to our life, health, and other fundamental values (Kang 2016). Even though security and safety terminologies seem different, the management of both types of circumstances is based on the same concepts, which are: i) discovering underlying events, ii) applying efficient procedures and plans to mitigate threats and to keep people and values safe from harm or injury, and iii) managing crisis and recovering from it. This research topic brings challenges for either cross-sectoral and thematic researchers and practitioners (Olsen, Kruke, and Hovden 2007).

As digitalisation continues to elaborate and expand in every area, the risks and threats facing society are evolving, and even more complicated, on a large scale. The advantages and convenience of digital are quick and straightforward, which are vital aspects to adapt to the modern appetite for real-time processing. Therefore, various spaces (e.g., email, SMS messaging, e-commerce, social networking service, and smart systems) can be targeted and intercepted by savvy hackers. These issues can significantly reduce our trust and increase insecurity to the same extent. It is precisely this urgency that requires a practical approach.

Artificial Intelligence (AI) opens a new perspective to maintain public security and safety by providing investigative assistance with a human-grade precision (Cath et al. 2018). Also, there is a critical need for automated solutions. For these reasons, targeted applications of AI to the domain of security and safety have recently come into concentration. This paper portrays AI-enabled use cases, which can be considered as opportunities to come up with pragmatic tools and solutions for helping address some current pressing challenges.

We introduced the background and emphasised our motivation in this section. The rest of this paper encompasses the following structure. In the next section, the necessary research methodologies will be given. Further, we will provide use cases on information validation and surveillance enhancement with the support of AI. Finally, we will draw essential conclusions and state future directions in the last section.

Research Methodology

Pressing Issues and Challenges

Several issues, which are not previously placed in the central, have now become the main focus. Examples involve a rising number of disinformation and insecurity incidents. They are becoming concurrently a premise for, and a threat to, societal security and safety.

Disinformation: Disinformation (i.e., false or misleading information) is generally not an emerging phenomenon; however, with the popularity of online platforms, it has become an increasingly sophisticated, deliberately circulated, and regularly utilised tool to achieve hostile targets and to cause harm. The spreading of disinformation poses an essential threat to societies and has adverse impacts on the quality of public life, stability, and societal security. For example, the outbreak of disinformation regarding COVID-19 has disseminated rapidly and widely across social networking services (Apuke and Omar 2020), endangering safety and impeding the recovery. Further, we are currently stepping into even more dedicated fake news. Not only text but also audio, photo, and video can be controlled and manipulated at will. In only 3.7 seconds, an algorithm named Deep Voice utilise snippets of voices to mimic the original one in order to create new speech, accents, and tones (Cole 2018). Augmented Reality (AR) and Virtual Reality (VR) will be the next-gen targets for disinformation with the upper realm of complexity and severe significance. Popular platforms (e.g., Facebook,

[*] Corresponding author

AAAI Fall 2020 Symposium on AI for Social Good.

Twitter, and YouTube) are concentrating on tackling online disinformation and limiting its circulation. Nevertheless, we still need to deepen our comprehending of the dangers of fake news and disinformation for well-informed and pragmatic societal security and safety planning. It is therefore necessary to research and develop advanced AI models that are able to identify fake news effectively and automatically.

Insecurity: The problem of insecurity and the feeling of insecurity are demanding immediate actions and efficient solutions. For this reason, a mass amount of cameras can be seen everywhere (e.g., on the streets and in businesses) in large cities. Law enforcement can place reliance on this footage to investigate crimes after the fact for prosecuting the guilty and catching criminals. Although surveillance cameras are inexpensive, the workforce necessary to keep track of and analyse them is expensive; hence, usually, videos from these cameras are only referred after critical events are known to have taken place. We find it unrealistic and infeasible for human observers to monitor and examine all the video streams with high accuracy. By leveraging AI-powered surveillance technologies, we enable the capacity to seek through more video more efficiently, to comprehend the full value of video surveillance, and to achieve expected results automatically while requiring less human intervention for video investigation.

Societal AI Research Cycle

This section introduces the action research cycle proposed by (McTaggart and Kemmis 1988) We follow the applied research method, which means the application of AI techniques into practice to address the risky situation of societal security and safety, conducted to solve real problems (i.e., use cases).

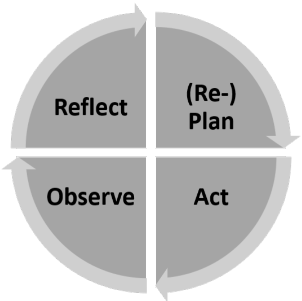

Figure 1: Action research cycle (self-reflective spiral (McTaggart and Kemmis 1988)).

According to (Kemmis, McTaggart, and Nixon 2013), action research is rarely as neat as this spiral of self-contained cycles of planning, acting and observing, and reflecting suggests in reality. Therefore, the process is likely to be more fluid, open, and responsive. In this regard, we repeat each cycle in a short period by following the agile methodology in computer engineering. The four steps of the action research cycle, which are depicted in Figure 1, are explained as follows.

- **Plan:** include problem definition, situation analysis, team vision, and strategic plan.
- **Action:** involve the implementation of the strategic plan.
- **Observation:** encompass monitoring and evaluation.
- **Reflection:** on the results of the evaluation.

Quantitative methods might always not be a correct evaluation of the AI techniques due to the characteristics of the societal security and safety domain. Therefore, We also need qualitative research methods in relevant use cases. For example, data collection for the study purpose is done by conducting interviews with organisation stakeholders, captivating opinions from industry experts, referring to existing literature, using principal consultants as a secondary source of information on initiatives adapted in similar organisations elsewhere to about with the trends. The study also analyses survey data available for stakeholders, relevant organisations, general observation, and end-users (including citizens) and an independent survey from IT professionals on the AI field. In the following section, two AI-enabled use cases for societal security and safety are described as preliminary studies based on literature review.

AI-enables Use Cases

Through use cases, we aim at investigating methodological, societal, technological issues, which in turn contribute to benefit from AI-based technologies, frameworks, and services.

Information Validation

Diverse thoughts and opinions (Long, Nghia, and Vuong 2014) are valued in modern society. Often it is called “cognitive diversity" and can counter group-think and enables better decision-making (Carey et al. 2016). Ironically, the cognitive diversity of a population is also being exploited in an entirely different way today. Instead of consolidating different perspectives and world-views into a superior consensus, new information technologies, such as online boutique news, social networks, and microblogs, take advantage of cognitive diversity by isolating subpopulations and catering to their idiosyncratic opinions. It often leads to giving people the illusion that they are in the ideological majority (Cybenko and Cybenko 2018). As such, cognitive diversity can be regarded as the Petri dish in which “fake news" thrives (Carey et al. 2016). Consequently, it is typically challenging to judge and accept the information that contradicts someone’s prior beliefs and world-views as truthful (Cybenko, Giani, and Thompson 2002).

As AI’s role in defeating cognitive safeguards, people now have a broader choice of information sources that they can self-select to align with whatever niche beliefs they may already have (Carey et al. 2016). It creates audiences with similar, idiosyncratic beliefs, and they can be identified and labelled using AI-based natural language and social network techniques (Hemavathi, Kavitha, and Ahmed 2017). After the audience identification, the content in the information can be adjusted to that audience. While human reporters and writers populate mainstream news and information sources, it is now possible to robotically generate news stories using AI-based software (WashPostPR 2016). Combining such technologies, we can imagine near-future AI-powered systems that will write a news article with minimal or no human intervention (Cybenko and Cybenko 2018). Besides, users self-select their sources and tend to see content consistent with their beliefs. And they then gain trust (Nguyen et al. 2017) in those sources. Once such community sources have been identified, AI technologies can author professional-looking websites with minimal human effort, catering to ideological niches (Tselentis 2017). Techniques for classifying news as “real” vs “fake” (or rumours vs non-rumours) generally fall into two categories. One class of methods uses linguistic and semantic analysis of written content to discriminate while the other uses dissemination patterns and rates to classify different types of news. Some approaches use both of them (Subrahmanian et al. 2016; Kwon, Cha, and Jung 2017).

Because the scale and scope of fake news claims will probably make human-based assessments about the veracity of information unsustainable (Alvarez 2018), identifying wrong information like “fake news" is a significant potential application of AI.

Surveillance Enhancement

Intelligent video surveillance based on AI is beneficial for monitoring of physical assets, large spaces, or significant events, for example, open-air concerts or film festivals. Since it is challenging to be in various places at once, we can rely on AI to detect violence or to analyse crowd behaviour for sending alerts if something is behaving abnormally. Beginning with a targeted video, we can apply object detection and identification to discover and locate unusual objects. The recognition can be categorised at either characteristic-based (Marcialis and Roli 2003) or behaviour-based (Robertson, Reid, and Brady 2008) level. Furthermore, we can train the AI models to determine potentially dangerous objects such as sharp objects, glass items, and weapons.

At the characteristic-based level, the analysis can be conducted by leveraging either face, head (Ishii et al. 2004), or body features. Given a single query video, or images extracted from this video, AI allows searching for the occurrence of a specific person. This gives us an opportunity to trace and discover his suspicious behaviours. In addition to that, we can estimate his gender, age (Antipov et al. 2017), and emotion (Jain, Shamsolmoali, and Sehdev 2019) as well. Apart from previous applications, AI-enabled surveillance enhancement still has other uses. By examining street footage, AI can determine vehicles concerning a set of attributes. For example, we can know exactly how many blue bus that passed through a specified location in a particular period. Where this becomes more helpful is when we want to find a stolen vehicle, and require a result promptly.

We aim at by analysing and detecting abnormal human actions at the behaviour-based level. The target is to anticipate whether a harmful event can occur because of an unusual behaviour (Ko and Sim 2018), even a few minutes in advance; for example, detecting abnormal driving (Huang et al. 2019) can help prevent an accident. The selection of techniques is influenced by two types of scene density that are un-crowded (i.e., single or a small number of people) and crowded. In un-crowded scenes, falling (for older adults), loitering (staying in a public location without apparent purpose for a long period), and violent actions (e.g., chasing and fighting) are useful to detect. On the other hand, it isn’t easy to monitor and analyse the behaviour of each person separately in crowded scenes. Possible approaches are crowd density estimation (i.e., assessing a crowd status), crowd motion detection (i.e., identifying behaviour pattern in a group), and crowd tracking (i.e., deriving trajectories of the movements).

Currently, deep learning algorithms and models (Zhou et al. 2016; Pérez-Hernández et al. 2020) are demonstrating their effectiveness in a large crowd at all crisis-related conditions, even in real-time (Pennisi, Bloisi, and Iocchi 2016; Nawaratne et al. 2019).

Conclusion

As digitalisation continues to elaborate and expand in the humanitarian domain, the risks and threats facing society are evolving, and even more complicated, on a large scale. In this paper we have illustrated AI-enabled use cases, which can be considered as opportunities, to come up with pragmatic tools, solutions, and service for addressing some current issues.

Besides, several challenges are needed to be taken into account. Human-level is the most important challenge in AI. We can develop a model with 80-90% accuracy; nonetheless, humans can achieve even absolute precision in all aforementioned use cases. Therefore, it is necessary to balance and keep humans on edge for AI systems and services. Data privacy is another critical challenge since AI-based algorithms learn from and make predictions based on data; many of them are personal and sensitive. This data can be in the target of bad purposes or of unlawful intents. Hence, we need to consider if or how to address the use of personal information in AI systems. We also need to seek appropriate methodologies to guarantee the protection of data while retaining the significant and potential benefits of big data analytics.

Acknowledgments

This work is supported by the INTPART BDEM project (grant no. 261685/H30).

References

Alvarez, E. 2018. Facebook’s approach to fighting fake news is half-hearted. https://www.engadget.com/2018/07/13/facebook-fake-news-half-hearted, accessed on 17.09.2020.

Antipov, G.; Baccouche, M.; Berrani, S.-A.; and Dugelay, J.-L. 2017. Effective training of convolutional neural networks for face-based gender and age prediction. *Pattern Recognition* 72: 15–26.

Apuke, O. D.; and Omar, B. 2020. Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telematics and Informatics* 101475.

Carey, J. M.; Nyhan, B.; Valentino, B.; and Liu, M. 2016. An inflated view of the facts? How preferences and predispositions shape conspiracy beliefs about the Deflategate scandal. *Research & Politics* 3(3): 1–9.

Cath, C.; Wachter, S.; Mittelstadt, B.; Taddeo, M.; and Floridi, L. 2018. Artificial intelligence and the 'good society': the US, EU, and UK approach. *Science and Engineering Ethics* 24(2): 505–528.

Cole, S. 2018. Deep Voice Software Can Clone Anyone's Voice With Just 3.7 Seconds of Audio. https://www.vice.com/en_us/article/3k7mgn/baidu-deep-voice-software-can-clone-anyones-voice-with-just-37-seconds-of-audio, accessed on 17.09.2020.

Cybenko, A. K.; and Cybenko, G. 2018. AI and fake news. *IEEE Intelligent Systems* 33(5): 1–5.

Cybenko, G.; Giani, A.; and Thompson, P. 2002. Cognitive hacking: A battle for the mind. *Computer* 35(8): 50–56.

Hemavathi, D.; Kavitha, M.; and Ahmed, N. B. 2017. Information extraction from social media: Clustering and labelling microblogs. In *Proceedings of the 2017 International Conference on IoT and Application (ICIOT), Nagapattinam, India, 19-20 May 2017*, 1–10. IEEE.

Huang, W.; Liu, X.; Luo, M.; Zhang, P.; Wang, W.; and Wang, J. 2019. Video-based abnormal driving behavior detection via deep learning fusions. *IEEE Access* 7: 64571–64582.

Ishii, Y.; Hongo, H.; Yamamoto, K.; and Niwa, Y. 2004. Face and head detection for a real-time surveillance system. In *Proceedings of the 17th International Conference on Pattern Recognition (ICPR), Cambridge, UK, 26-26 August 2004*, 298–301. IEEE.

Jain, D. K.; Shamsolmoali, P.; and Sehdev, P. 2019. Extended deep neural network for facial emotion recognition. *Pattern Recognition Letters* 120: 69–74.

Kang, H.-J. 2016. A Study on Analysis of Intelligent Video Surveillance Systems for Societal Security. *Journal of Digital Contents Society* 17(4): 273–278.

Kemmis, S.; McTaggart, R.; and Nixon, R. 2013. *The action research planner: Doing critical participatory action research*. Singapore: Springer Science & Business Media.

Ko, K.-E.; and Sim, K.-B. 2018. Deep convolutional framework for abnormal behavior detection in a smart surveillance system. *Engineering Applications of Artificial Intelligence* 67: 226–234.

Kwon, S.; Cha, M.; and Jung, K. 2017. Rumor detection over varying time windows. *PloS One* 12(1): e0168344.

Long, N. H.; Nghia, P. H. T.; and Vuong, N. M. 2014. Opinion spam recognition method for online reviews using ontological features. *Tạp chí Khoa học* (61): 44–59.

Marcialis, G. L.; and Roli, F. 2003. Fusion of face recognition algorithms for video-based surveillance systems. In Foresti, G. L.; Regazzoni, C. S.; and Varshney, P. K., eds., *Multisensor surveillance systems: the fusion perspective*, 235–249. Boston, MA, USA: Springer.

McTaggart, R.; and Kemmis, S. 1988. *The action research planner*. Melbourne, Victoria, Australia: Deakin university.

Nawaratne, R.; Alahakoon, D.; De Silva, D.; and Yu, X. 2019. Spatiotemporal anomaly detection using deep learning for real-time video surveillance. *IEEE Transactions on Industrial Informatics* 16(1): 393–402.

Nguyen, H. L.; Lee, O.-J.; Jung, J. E.; Park, J.; Um, T.-W.; and Lee, H.-W. 2017. Event-driven trust refreshment on ambient services. *IEEE Access* 5: 4664–4670.

Olsen, O. E.; Kruke, B. I.; and Hovden, J. 2007. Societal safety: Concept, borders and dilemmas. *Journal of contingencies and crisis management* 15(2): 69–79.

Pennisi, A.; Bloisi, D. D.; and Iocchi, L. 2016. Online real-time crowd behavior detection in video sequences. *Computer Vision and Image Understanding* 144: 166–176.

Pérez-Hernández, F.; Tabik, S.; Lamas, A.; Olmos, R.; Fujita, H.; and Herrera, F. 2020. Object detection binary classifiers methodology based on deep learning to identify small objects handled similarly: Application in video surveillance. *Knowledge-Based Systems* 194: 105590.

Robertson, N.; Reid, I.; and Brady, M. 2008. Automatic human behaviour recognition and explanation for CCTV video surveillance. *Security Journal* 21(3): 173–188.

Subrahmanian, V.; Azaria, A.; Durst, S.; Kagan, V.; Galstyan, A.; Lerman, K.; Zhu, L.; Ferrara, E.; Flammini, A.; and Menczer, F. 2016. The DARPA Twitter bot challenge. *Computer* 49(6): 38–46.

Tselentis, J. 2017. When websites design themselves. https://www.wired.com/story/when-websites-design-themselves, accessed on 17.09.2020.

WashPostPR. 2016. The Washington Post experiments with automated storytelling to help power 2016 Rio Olympics coverage. https://www.washingtonpost.com/pr/wp/2016/08/05/the-washington-post-experiments-with-automated-storytelling-to-help-power-2016-rio-olympics-coverage, accessed on 17.09.2020.

Zhou, S.; Shen, W.; Zeng, D.; Fang, M.; Wei, Y.; and Zhang, Z. 2016. Spatial–temporal convolutional neural networks for anomaly detection and localization in crowded scenes. *Signal Processing: Image Communication* 47: 358–368.