DOI: 10.38146/BSZ-AJIA.2024.v72.i7.pp1257-1273

The Intelligent Threat: How Artificial Intelligence Can Compromise Our Security

Levente Tóth
PhD, operational director, assistant lecturer
TVT Security Ltd.
Ludovika University of Public Service,
Faculty of Law Enforcement
toth.levente@uni-nke.hu

Abstract

Aim: The aim of the study is to draw attention to the dangers of using artificial intelligence.

Methodology: Alongside a relevant literature review, the author illustrates the aspects of artificial intelligence jeopardising our security by providing examples and addresses the existing and evolving regulatory environment.

Findings: Artificial intelligence can directly or indirectly pose a threat to our security. The risks associated with artificial intelligence, coupled with the current rapid technological advancement, make it imperative to establish appropriate and adaptive continuous regulations to ensure the increasing use of AI comes with minimal negative consequences.

Value: The study explores previously overlooked features that compromise security. Its findings can contribute to understanding how artificial intelligence can endanger our security on both narrower and broader societal levels.

Keywords: artificial intelligence, morphing, smart home, facial recognition

Introduction

Artificial intelligence (hereafter: AI) is one of the most decisive areas of technological development today. For a long time now, it has become not only

English-language republication. The Hungarian version of this article was published in Belügyi Szemle 2024, issue 7. DOI link: https://doi.org/10.38146/BSZ-AJIA.2024.v72.i7.pp1225-1239

a scientific experiment, but a determining factor in the everyday life of all of us. AI has enormous benefits that span home use, the corporate sector and a wide range of industries. Tasks that previously took long hours or involved a high rate of human error can now be done quickly and accurately with the help of machine processes. The dynamic development and increasingly widespread application of AI technology plays a key role in the transformation of our modern society and economy. However, we should not ignore the factors associated with the application of AI, which indirectly and directly threaten the safety of human existence, which are discussed much less often.

What is AI?

AI is a field of computer science that aims to model and simulate the mechanisms of human thinking and decision-making using machine systems. The goal of AI is to imitate intelligent behaviour similar to human activity, including learning, problem-solving, decision-making, visual perception, and speech recognition (Russel & Norvig, 2010). Machine learning is a subset of AI that develops algorithms that 'learn' from data, and their performance increases as the data is learned. Machine learning algorithms cannot be programmed to perform specific tasks, but learn how to perform tasks through data-based training, i.e. they learn implicit rules from many examples in the database. That is, data is used to make informed estimates about the future. From a law enforcement perspective, a good example is so-called 'predictive policing', where AI is used, for example, to assess the predicted risk of crime (URL1). The Dutch police, for example, have a similar system. The so-called Crime Anticipation System processes crime patterns and predicts where and when crimes are likely to occur. This allows monitoring and prevention activities to be adjusted based on predicted risks (Oosterloo & van Schie, 2018).

Deep learning is a field of machine learning based on artificial neural networks. A neural network is a computer model inspired by the biological neural network of the human brain. Neural networks usually consist of interconnected neurons. One of the most important elements of deep learning is the use of multilayer neural networks. These networks contain multiple layers that build on each other, where each layer performs different mathematical operations on the input data. During these layers, the data forms higher and higher level abstractions, which enables the system to solve complex tasks. These include image recognition, speech recognition, natural language processing, control of autonomous vehicles, help with diagnosis in medicine, and make quality control more efficient in industrial production. In addition to these, they also play a key role in face recognition technologies, which are increasingly gaining ground not only in security systems, but also in user software (for example, automatic image tagging). Such machine learning models are constantly evolving, and as they learn on more and more data, they become more and more efficient and reliable in performing specific tasks.

In recent years, AI has advanced to the point where AI systems can recognise objects and texts, make real-time decisions, and mimic human interaction at some level. They are also capable of understanding human speech patterns and responding adequately, opening new dimensions in many areas of human life. Algorithms and machine learning methods using AI can predict consumer preferences and market trends. Based on the analysis of user profiles and behaviour patterns, AI can provide relevant and interesting recommendations and content to individuals. This provides a significant advantage in areas ranging from online commerce to music listening and content delivery. Furthermore, AI enables quick decision-making. Large amounts of data can be analysed and interpreted in a short time, which is crucial in corporate decision-making and crisis management.

The possibility of choice

In the spring of 2023, the *New York Times* published a remarkable article entitled *You Can Have the Blue Pill or the Red Pill, and We're Out*, in which the authors expressed their views on the dangers and possibilities of AI (URL2). The title of the article refers to the cult film The Matrix, in which the main character, Neo, must make a critical decision: choose the blue pill and stay in the familiar but illusory world, or choose the red one and face the painful truth. The article draws a parallel between this choice and the two perspectives on AI. The blue pill represents those who see AI as just a useful tool without much need for critical thinking or caution. The red pill, on the other hand, represents those who acknowledge the potential dangers of AI and are prepared to take action to address them. The authors argue that AI is not merely another technological advancement but rather one of the greatest challenges and opportunities in human history. They believe that AI has the potential to significantly improve various aspects of human life, from healthcare to education to sustainability. The article encourages individuals to take the red pill but also emphasises the importance of being well-informed, critical, and proactive in addressing the challenges and opportunities presented by AI. The central question posed is whether society is prepared to confront the hidden challenges of AI and harness its potential or whether it will choose to remain in comfortable ignorance and allow others to make decisions on its behalf. The future of humanity may depend on this decision.

The development of AI has started a new era of the transformation of society, the consequences of which we cannot yet fully foresee. What would it mean for humans to live in a world where a large percentage of blog posts, stories, newspaper articles, books, images, tunes, laws, and devices are created by non-human intelligence, and this AI is superhumanly efficient at exploiting the weaknesses, biases, and addictions of the human mind, while he would even be able to form an intimate relationship with the person. In such a world, the impact would be completely unpredictable. In certain fields, such as strategy games (Chess, Go, Shogi), AI has long surpassed human capabilities (URL3). However, this development is not only limited to games, but can also extend to other areas such as art, politics, or religion. This can potentially lead to changes in power dynamics and decision-making processes, as well as affect the structure of human society. In the field of art, non-human intelligence can create works that deeply touch human emotions, and can even compete with human-made art, possibly surpassing it. In the field of politics, with the help of the so-called Deep Fake technology, it is possible to spread disinformation aimed at humiliating, harassing, blackmailing, or destroying the image and credibility of a selected individual. In addition, it can be used to destabilise market players, thereby disrupting financial markets, and even inciting social discontent. Similarly, in the realm of religion, non-human intelligence can reshape spiritual narratives and practices, potentially changing belief systems on a global scale. Also among the dangers is the potential loss of human autonomy and self-determination. It is therefore extremely important to shape our collective cultural, social and ethical frameworks in such a way that they support the preservation of human autonomy and self-determination in this new environment. Social dialogue and cooperation will be of paramount importance when dealing with such changes.

From the above, it is clear that in the world of non-human intelligence, which exerts a significant influence on human life, many important questions arise, which prompt us to consider them from an ethical, philosophical and practical point of view. How would humanity adapt to such a paradigm shift, and how would the relationship between human and non-human intelligence evolve?

The issue of independent decision-making and responsibility in the field of AI systems is also a serious challenge. AI systems are becoming increasingly complex and able to perform tasks that were previously exclusively human activities. As a result, the question arises as to who bears the responsibility for any errors or damages that may occur as a result of the decisions made by AI. A serious concern is that the processes and logic behind the decisions made by AI are often difficult to understand and control. As a result, it can be a challenge for people to fully control the decisions made by AI and to delimit responsibility.

The dangers of AI are further increased by the issue of data protection and privacy. The operation of AI systems requires a huge amount of data, which can often include personal data. Due to the amount and sensitivity of data collected during the development of AI technologies, special attention must be paid to the protection of personal data. Data protection regulations such as the EU General Data Protection Regulation (GDPR) must be complied with when designing and using AI systems. Due to the risks of unauthorised access, data misuse, or data theft, it is important that AI systems are designed to respect user privacy and ensure data security. This includes secure data storage, restricting access and informing users about how and for what purpose their data is used.

Another important aspect of the dangers of AI can be the increase of social inequalities. The operation and decision-making of AI systems largely depends on the data and samples with which they are taught and trained. If this data is distorted or discriminatory, then the decisions made by the AI may also be distorted and discriminatory, and the so-called algorithmic discrimination[1] may develop. And this can contribute to the growth of social inequalities.

The importance of responding to and managing AI threats is increasing as technology advances. It is important to consider these sources of danger during the development and application of AI and to find solutions that minimise these risks. Adhering to ethical guidelines and increasing transparency are crucial for the responsible and sustainable development of AI. At the same time, taking advantage of the potential benefits of AI, we must act with due caution and responsibility in order to ensure that the development of AI is beneficial to humanity and does not endanger human values and rights.

In addition to the dangers described above, AI also has many direct security-threatening uses that threaten our personal safety and ‘only’ have an indirect effect on human communities.

[1] Algorithmic discrimination means when, as a result of an automatic decision system, individuals or groups of people are treated differently based on certain characteristics (such as skin color, gender, age, religion, sexual orientation, etc.).

Smart homes

Smart homes, represent the convergence of modern technology and the domestic environment, where by connecting different devices and systems, they increase the comfort, safety and efficiency of energy management for users. The central unit of smart homes is an intelligent control system, which can be a dedicated local central server or a widely available cloud platform. Smart home systems can include automated control of lighting, heating, air conditioning, security cameras, locks, alarms, consumer electronics and home appliances. These devices often communicate with each other and the central controller or the cloud service via wireless technologies, allowing the user to control their home remotely. The intelligence of smart homes lies not in the fact that they enable remote control, but in their ability to learn the habits and preferences of their users. For example, a smart thermostat can 'observe' the daily routine of residents and adjust the temperature based on when they are home and when they leave. Smart lighting systems are able to adapt to natural light conditions, thereby reducing energy consumption and increasing the feeling of comfort. In terms of security, smart homes offer features such as cameras connected to motion sensors that can send an alert to the owner if unusual activity is detected. Smart locks allow keyless entry and this can be remotely monitored or controlled, for example allowing maintenance staff to have supervised access to certain areas. In the field of energy management, smart devices help to optimise energy use and reduce the electricity bill. For example, smart appliances are able to switch on only when energy prices are lower, or they can take advantage of renewable energy sources such as solar panels. In addition to all this, smart homes also play an important role in the lives of the elderly and disabled, providing them with the opportunity to live independently at home without help. With the help of voice control and automated routines, they can simplify their daily activities and improve their quality of life. Overall, smart homes represent the future of home living, where smart technology helps increase energy efficiency, enhance security, and generally improve the quality of life. However, in the case of inadequate protection, we can pay a heavy price for the comfort achieved in this way. Since smart home systems are connected to an IT network, there is a risk of external hacker attacks. Smart home devices can collect and transmit various data to the cloud or to other devices. This data may contain personal information that unauthorised persons may attempt to obtain or use. The list of dangers does not end here.

For years, the development of AI has made it possible to control various devices with voice commands. This so-called speech recognition is not an easy task. Speech recognition is the area of AI that deals with the detection, analysis and interpretation of spoken human language. This includes distinguishing words and sentences and converting them into text (speech-to-text), or vice versa, converting text into speech. Devices that can be activated by voice are becoming more and more common every day. There are more than 300 million smart speakers worldwide (URL4). With voice assistance built into smartphones, this number rises to over four billion. Popular platforms such as Google Assistant, Apple Siri, Amazon Alexa, Samsung Bixby, or the systems of China's Baidu, Alibaba, and Tencent have a significant market share. Digital assistants are now like a digital personal assistant that reminds you of daily tasks, provides weather forecasts, and provides personalised information. All the user has to do is give a clear verbal command. Speech recognition technology converts speech into text, then natural language processing interprets the written information and determines what action is required. While humans were able to speak and interpret speech even before the invention of writing, computers find it easier to process written language than the spoken word. Speech recognition is a significantly more complex task due to the variability of spoken language and possible noise in audio streams. Selecting, identifying and converting words into the type of text that a computer can process is extremely challenging. When we speak, the sounds we make are not separated into separate words. What a computer perceives is very similar to what a person hears when listening to a language that is completely unknown to them. In this case, we hear almost a continuous stream of sounds, in which it is very difficult to distinguish individual words. The problem of speech recognition is thus fundamentally different from the interpretation of written language or images. Unlike machine vision and natural language processing, speech recognition involves the processing of a single input variable—sound waves—that changes dynamically over time. The big challenge is distinguishing the words and sentences within this input so that they can be translated into a language that the algorithm can process. Another challenge is that part of the meaning of speech is conveyed by changes in volume, rhythm, and timbre, which are characteristics of spoken language. Effective interpretation is therefore more than simply distinguishing words. Phonetic aspects must also be explored and interpreted in order to determine the meaning of what is being said. Another pitfall is homophones, which are words that sound the same but mean different things, such as eye, I, or no, know. Their interpretation depends on the context. Both from the narrow context of the sentence and from the broader context of the situation. As in other fields, the development of machine learning has led to progress in the field of speech recognition, as it has become possible to process much larger amounts of speech data for training algorithms. Relatively successful practical applications of speech-to-text and text-to-speech conversion are now viable, provided that the speech is accurate in terms of both hearing and content. However, verbal communication full of slang is often not clear to AI. Consequently, speech recognition technology has not yet reached a level where it can be reliably used widely and for all purposes. However, properly articulated sentences and instructions are excellently interpreted by the AI. In order to increase comfort, these devices that receive and interpret voice commands can be connected to the central controls of smart homes, and by saying appropriate command words, various devices can be activated, switched on or scenes can be run. However, there are many attack opportunities against such systems.

In recent years, many experiments carried out by researchers have clearly proven that it is possible to send hidden commands imperceptible to the human ear to various voice-based control systems. One such attack method involves transposing control commands into the ultrasound range, so that the commands are inaudible to the human ear, but can be detected by the microphones used by voice control systems in smart homes. (Zhang et al., 2017). During the experiment, the researchers were able to secretly activate and control the voice assistant systems through smartphones and smart speakers. As a result, the devices dialled phone numbers or opened web pages without the human ear detecting this process. Such technical possibilities can pose a serious risk in the hands of persons with malicious intent, as they can even be used to transfer money or make online purchases. In addition, when connected to smart homes, different types of commands can be issued, such as turning lights on or off, opening and closing doors, changing the temperature of the thermostat, or bypassing other security systems.

The IT experts of the Horst Görtz IT Security Institute in Bochum used a completely different method (Schönherr, Kohls, Zeiler, & Holz, 2018). In their experiments, they use a psychoacoustic model of human hearing. While the ear and brain are busy processing a sound at a particular frequency, it is unable to hear other low-volume sounds for a few milliseconds. In these parts, the researchers hide the secret commands of the machines, which additional information appears to the human ear as random static noise. The attack method is based on the so-called adversarial example. Adversarial examples are special inputs that are created to confuse a neural network, resulting in the misclassification of a given pattern. These manipulated input patterns are imperceptible to the human ear or eye, but they affect the operation of the machine learning model and produce false predictions. In the case of images, we can also say that they are optical illusions created for machine learning. This brings us to image manipulation techniques.

Facial recognition systems

Facial recognition is an AI-based computer technology used to find and identify human faces in digital photographs or live or recorded video images. Facial recognition technology is often used to monitor and track people in real time. As a first step, the face recognition system starts the process by detecting faces. For this, we can use different algorithms, such as Viola-Jones,[2] or other face detection algorithms based on deep learning. In facial recognition, algorithms identify points on faces that serve as unique features. These points are usually called facial recognition points, with the help of which the algorithms are able to identify. After the face detection is done, the system can perform further processing steps, such as determining the contours of faces or segmenting their area. During the extraction of various features (such as facial features, eye location, etc.), the system converts the face into numerical values that are used for later comparison. The facial recognition system compares the extracted features with the features stored in the database to determine which one corresponds to the recognised face. Finally, the system makes a decision about which of the stored reference faces the recognised face is most similar to.

China leads the world in government use of facial recognition systems. The technology is widely used by the police to monitor urban public areas. According to Surfshark's 2019 survey of 194 countries, 109 countries in the world use or have approved the use of facial recognition technology for surveillance purposes, including 32 countries in Europe (URL5). Compared to traditional non-biological recognition and physiological feature recognition technology, face recognition technology has special technical advantages. The attributes and features of a facial image are sufficient to establish a person's identity. The facial recognition technology can also extract additional information suitable for personal identification, such as age, gender and race, from the images. Some Chinese camera manufacturers make it possible, for example, to filter the Uyghurs living in large numbers in the western part of China, in the Xinjiang province, according to their race, thanks to the facial recognition algorithm integrated into their cameras (URL6). AI can be used not only in the field of facial recognition, but also when applied in other video analytics software to warn when abnormal actions occur (Tóth, 2018). Taking advantage of this, another Chinese

[2] The Viola-Jones algorithm is named after two computer vision researchers (Paul Viola and Michael Jones) who proposed the method in 2001 in their paper Fast object detection using an augmented cascade of simple functions (Viola & Jones, 2001). Despite being an outdated framework, its application has proven exceptionally remarkable in real-time face recognition (although it can be applied to other object classification as well). This algorithm is slow to train, but it can detect faces in real time at high speed.

manufacturer's AI-based analytics includes nine different types of alerts related to mass movement. These include 'gathering of crowds to disturb public order', 'unlawful assembly, march, demonstration', threatening to 'petition', and 'religious' and the Chinese spiritual movement known as 'Falun Gong' (URL7). Such discrimination against minority communities or ethnicities raises serious ethical and moral questions.

Due to the possibility of linking information, there is a growing concern about the harmful effects of video surveillance on our privacy. The main problem is the profile created from the various available information that can be used to invade privacy. By incorporating an automatic facial recognition and analysis algorithm and connecting different databases or network systems, significant information can be extracted. Legislators find it difficult to keep up with the rapid pace of technological development. The European Union's draft law on AI (URL8) proposes to limit the use of facial recognition systems in public, and the European Parliament has urged a ban on the technology (URL9). In July 2021, the European Data Protection Authority and the European Data Protection Commissioner adopted a joint opinion on the European Commission's proposal for a regulation establishing harmonised rules[3] for AI.[4] In it *„ the EDPB and the EDPS call for a general ban on any use of AI for anautomated recognition of human features in publicly accessible spaces – such as of faces but also of gait, fingerprints, DNA, voice, keystrokes and other biometric or behavioural signals – in any context."* (URL10).

Facial recognition systems are not infallible either. In an experiment by the McAfee Advanced Threat Research (ATR) team, it was proven that manipulated faces created with the help of AI can mislead facial recognition systems (URL11). Researchers have developed a method to convert the faces of two different people A and B into a fake face C. This fake, so-called morphed face C looks deceptively similar to person A to the human eye, but the AI-based face recognition system mistakenly identifies it as person B.

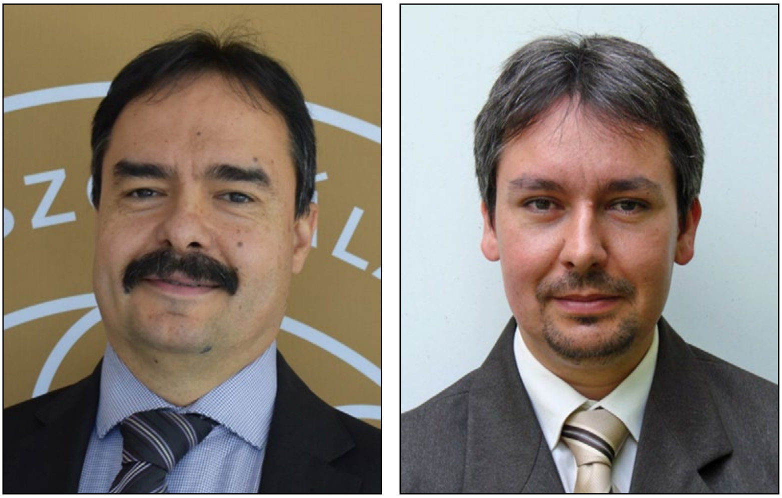
The basis of the method is the so-called Cycle-Consistent Generative Adversarial Network, i.e. CycleGAN for short. CycleGAN is a type of generative adversarial network (GAN) that can restyle images without requiring paired data. This machine learning model is used to transform between two different datasets. CycleGAN has two generators and two discriminators. Generators generate images from one dataset to another, and discriminators try to distinguish between

[3] The Artificial Intelligence Act (AI Act), codified regulation COM/2021/206 proposed by the European Commission, submitted on 21 April 2021.

[4] Following the finalisation of the manuscript, on 21 May 2024 the Council of the European Union (EU) announced the final approval of the landmark EU Artificial Intelligence Act. (Editorship.)

real and generated images. The essence of how CycleGAN works is that the generators are trained to perform the transformation between the two datasets so that the reconstructed images match the original images. In other words, the model is able to capture the features of the target domain and generate new images from the source domain that have the same features. CycleGAN is trained with a cycle-consistent loss that encourages the model to generate images that are indistinguishable from real images of the target domain. CycleGAN is a very useful tool for applications such as copying art styles, generating photorealistic images, or even synthesising maps from satellite images or aerial photographs.

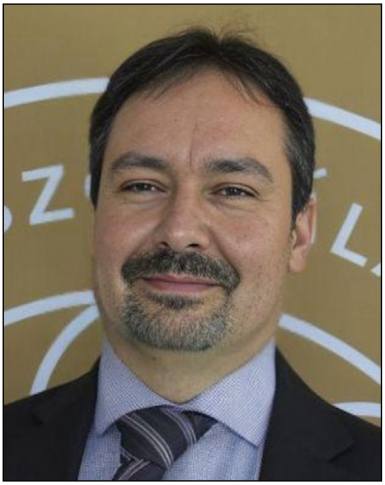
Figure 1
Original face A and B



Note. Figure prepared by the author.

Returning to our starting point, in the case of faces, the algorithm works by first generating a transitional face type AB (figure 2) from the original face A and B (figure 1).

Figure 2

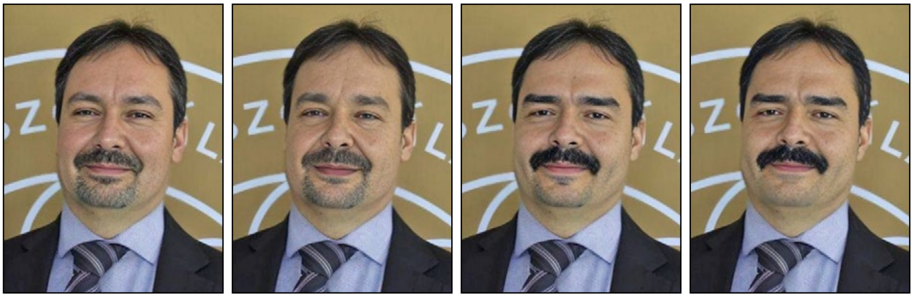


Transitional AB face
Note. Figure prepared by the author.

This temporary face type is then regenerated several times by the MI (Figure 3). This process continues until a face type is created that is deceptively similar to the original person A to the human eye, but the AI-based face recogniser clearly recognises the style features of person B in it.

To prove this, McAfee researchers tested the manipulated facial image they generated on a facial recognition system that normally performed the identification with 99% accuracy. However, in the case of manipulated faces, this system can be wrong with a probability of up to 95%, i.e. it determined with 95% certainty that person A is person B.

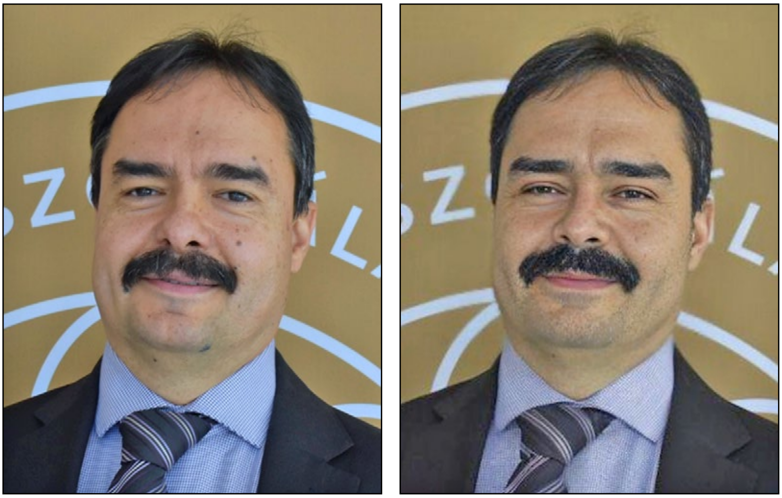
Figure 3
Progressive regeneration of faces



Note. Figure prepared by the author.

In order to achieve the perfect effect, during the training, a series of 1,500 images of faces were used, which were still images taken from video films made about persons. These images enriched the training data and ensured that the algorithm worked with a variety of facial expressions.

Figure 4
On the left is the original face A, on the right is the morphed face C



Note. Figure prepared by the author.

This type of image manipulation can pose a huge threat to identification systems based on AI-based facial recognition technology. The generated fake faces can be used to gain unauthorised access to a place or service, or to circumvent the automatic border control system with a fake passport of a wanted person, since the machine manipulation is invisible to the human eye without closer observation (Figure 4).

The fear of this kind of abuse justified the fact that in the summer of 2020, the law was proposed in Germany that would prohibit the use of morphed images in passports (URL12). Finally, after a long delay, the section of the Law on Strengthening the Security of Passport, Identity and Immigration Legal Documents will enter into force only on 1 May 2025, from which only a digital photograph can be presented to the authorities and offices in order to issue an identity card and passport (URL13). These pictures can only be taken at the issuing authority or at a registered and verified photographer, where secure image transmission via the cloud to the authority can be arranged.

In 2020, the first false arrest occurred in America thanks to the facial recognition system (URL14). Detroit police arrested Robert Williams outside his home on suspicion of stealing several watches from the local Shinola store. The suspicion was based on a video recording of a black man stealing from the store, which the police investigated with the help of facial recognition software. The software incorrectly identified the perpetrator as Williams. These cases point out that facial recognition systems cannot be considered infallible either, and it is important to conduct further research in order to make these technologies more reliable.

Facial recognition is by no means the only application of machine vision. It is also crucial for self-driving vehicles. Autonomous and semi-autonomous cars currently under development by Tesla, BMW, Volvo and Audi are equipped with multiple cameras that continuously scan the surrounding space and recognise objects, road signs, and traffic signs with the help of AI-based image classifiers and lamps. During an experiment, researchers at Kyushu University in Japan came to the startling conclusion that AI-based image classifiers can be fooled 73.8% of the time by changing just one pixel in the image. If the number of replaced pixels was increased to five, the success rate rose to 87.3%. This image manipulation is imperceptible to the human eye, but it can significantly distort the output produced by the AI. So, for example, AI interprets a picture of a frog as a ship or an airplane as a dog (Su, Vargas, & Sakurai, 2019). With this type of manipulation, it has already been possible to influence the traffic sign recognition function of self-driving cars and make the vehicle exceed the permitted speed (URL15).

Summary

There is still much potential and untapped potential for MI. How much of this will be positive and how much negative is yet to be seen. Various AI spoofing attacks, algorithmic discrimination, accidents involving self-driving cars, and dehumanisation associated with over-reliance on technology are good examples of what threatens the society of the future. The risks posed by AI and the current speed of technology development make appropriate and adaptive ongoing regulation even more urgent if the increasing use of AI is to have as few undesirable consequences as possible. Before his death, Professor Stephen Hawking warned that *„Success in creating AI could be the biggest event in the history of our civilisation. But it could also be the last – unless we learn how to avoid the risks. The rise of powerful AI will either be the best or the worst thing ever to happen to humanity. We do not yet know which.”* (URL16).

References

Oosterloo, S. & van Schie, G. (2018). The Politics and Biases of the "Crime Anticipation System" of the Dutch Police. *CEUR Workshop Proceedings*, *2103*, 30–41.

Russel, S. & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach* (3rd Ed.). Prentice Hall.

Schönherr, L., Kohls, K., Zeiler, S. & Holz, T. (2018). Adversarial Attacks Against Automatic Speech Recognition Systems via Psychoacoustic Hiding. *Network and Distributed Systems Security (NDSS) Symposium 2019 24-27 February 2019, San Diego.* https://doi.org/10.14722/ndss.2019.23288

Su, J., Vargas, V. D. & Sakurai, K. (2019). One Pixel Attack for Fooling Deep Neural Networks. *IEEE Transactions on Evolutionary Computation*, *23*(5), 828–841. https://www.theguardian.com/science/2016/oct/

Tóth A. (2018). Az élőerő munkáját segítő technikai megoldások. *Hadmérnök, 13*(2), 29–36.

Viola, P. & Jones, M. (2001). Rapid Object Detection using a Boosted Cascade of Simple. In Jacobs, A. & Baldwin, T. (Szerk.), *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 511–518). Computer Society. https://doi.org/10.1109/CVPR.2001.990517

Zhang, G., Yan, C., Ji, X., Zhang, T., Zhang, T. & Xu, W. (2017). DolphinAttack: Inaudible Voice Commands. *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (pp. 103–117). https://doi.org/10.1145/3133956.3134052

Online links in the article

URL1: *From Crime Mapping to crime forecasting: The evolution of olace-based policing. National Institute of Justice.* https://nij.ojp.gov/topics/articles/crime-mapping-crime-forecasting-evolution-place-based-policing

URL2: *You Can Have the Blue Pill or the Red Pill, and We're Out of Blue Pills.* https://www.nytimes.com/2023/03/24/opinion/yuval-harari-ai-chatgpt.html

URL3: *Game-playing DeepMind AI can beat top humans at chess, Go and poker.* https://www.newscientist.com/article/2402645-game-playing-deepmind-ai-can-beat-top-humans-at-chess-go-and-poker/

URL4: *Installed base of smart speakers worldwide in 2020 and 2024.* https://www.statista.com/statistics/878650/worldwide-smart-speaker-installed-base-by-country/

URL5: *The Facial Recognition World Map.* https://surfshark.com/facial-recognition-map

URL6: *As China Tracked Muslims, Alibaba Showed Customers How They Could, Too.* https://www.nytimes.com/2020/12/16/technology/alibaba-china-facial-recognition-uighurs.html

URL7: *Police in China can track protests by enabling 'alarms' on Hikvision software.* https://www.theguardian.com/world/2022/dec/29/china-surveillance-protests-alarms-cameras-hikvision

URL8: *Proposal For A Regulation Of The European Parliament And Of The Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts.* https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206

URL9: *REPORT on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters.* https://www.europarl.europa.eu/doceo/document/A-9-2021-0232_EN.html#title1

URL10: *EDPB-EDPS Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).* https://edpb.europa.eu/our-work-tools/our-documents/edpbedps-joint-opinion/edpb-edps-joint-opinion-52021-proposal_en

URL11: *Dopple-ganging up on Facial Recognition Systems.* https://www.mcafee.com/blogs/other-blogs/mcafee-labs/dopple-ganging-up-on-facial-recognition-systems/

URL12: *Germany bans digital doppelganger passport photos.* https://www.reuters.com/article/idUSKBN23A1YM/

URL13: *Gesetz zur Stärkung der Sicherheit im Pass-, Ausweis- und ausländerrechtlichen Dokumentenwesen.* https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&start=//*[@attr_id=%27bgbl120s2744.pdf%27]#__bgbl__%2F%2F*%5B%40attr_id%3D%27bgbl120s2744.pdf%27%5D__1704650921418

URL14: *This may be America's first known wrongful arrest involving facial recognition.* https://edition.cnn.com/2020/06/24/tech/aclu-mistaken-facial-recognition/index.html

URL15: *Hackers can trick a Tesla into accelerating by 50 miles per hour.* https://www.technologyreview.com/2020/02/19/868188/hackers-can-trick-a-tesla-into-accelerating-by-50-miles-per-hour/

URL16: *Stephen Hawking: AI will be 'either best or worst thing' for humanity.* https://www.theguardian.com/science/2016/oct/19/stephen-hawking-ai-best-or-worst-thing-for-humanity-cambridge

Reference of the article according to APA regulation

Tóth L. (2024). Az intelligens fenyegetés. Hogyan veszélyeztetheti a mesterséges intelligencia a biztonságunkat? *Belügyi Szemle*, *72*(7), 1257–1273. https://doi.org/10.38146/BSZ-AJIA.2024.v72.i7.pp1257-1273

Statements

Conflict of interest

The author has declared no conflict of interest.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Ethics

Corresponding author

The corresponding author of this article is Levente Tóth, who can be contacted at toth.levente@uni-nke.hu.