Artificial Intelligence:

Advances, Controversies, and Solutions — Convey Human conscience to AI

210016568

University of St Andrews

CS5010 - Artificial Intelligence Principle

Academic Year 2021/2, Semester 1 - Coursework 1

Lecturer: Ognjen Arandelovic

Word Count: 5366 words

Upload date: October 5, 2021

Abstract

Since the birth of the discipline of artificial intelligence at the Dartmouth Conference in 1956, the technical theory has continued to develop, and the application areas have been extended. Especially in recent years, the artificial intelligence in the field of computer vision has been developing rapidly due to the continuous improvement of computer computing power and the continuous development of sensor technology. In this paper, we start from two main areas of AI applications in computer vision which is medical image and face recognition. The paper will discuss the advances, controversies, and solutions to those problems for AI applications in medical image and face recognition from 2015 to 2021.

Key words: Artificial Intelligence, AI, Medical Image, Face Recognition, Consent, Discrimination, Deepfake

I. Introduction

According to Blasch et al. (2021), advances in sensor technology have made it easier to collect data especially image for machine learning and deep learning. These two are the main approaches to artificial intelligence, which has further led to the creation of data-driven models. At the same time, the significant increase in GPU arithmetic power has made the training and iteration of models significantly faster (Deloitte China 2020). These two reasons have led to tremendous advances and extraordinary achievements in artificial intelligence in areas such as image classification, autonomous driving, as well as semantic segmentation (Cui et al. 2019). However, the impact of artificial intelligence involves not only the field of computer science, but also many social issues that arise from its application scenarios (Strümke et al. 2021). Because the speed and extent of its own technological development exceeds the speed of the public's understanding of AI, and the scope of existing laws governing it. Furthermore, artificial intelligence is controversial in many of its application scenarios. For instance, data sources for artificial intelligence algorithm training may be illegally collected, and attacks against artificial intelligence algorithms in the areas of face recognition and autonomous cars driving can cause significant property damage and even life-threatening injuries to users.

This essay will employ the definition of artificial intelligence from Russell & Norving (2014), which is described as an intelligent system that can think and act rationally like a human being. In artificial intelligence, image data are easier to collect than other types of data, and this, as a result, has allowed various computer vision algorithms based on artificial intelligence such as convolutional neural network to develop rapidly. Therefore, this essay will focus on several artificial intelligence applications related to computer vision and put forward the argument that those applications did create advances from 2015 to present. Although it introduced certain controversies, these controversies can be addressed strategically.

This essay is divided into three main sections. These three sections will illustrate the advances artificial intelligence made from 2015 to present, the controversies and issues artificial intelligence brought, then will discuss the solution, which this essay contents to add morality and human consciousness to artificial intelligence.

II. Advances of Artificial Intelligence (2015 – 2021)

This section will discuss the advances of artificial intelligence in terms of two respects of the most popular AI applications in recent years, namely medical image and face recognition.

i. Medical Image

Medical imaging has gained significant momentum in a period of rapidly improving artificial intelligence devices and algorithms. Razzak et al. (2018) presented that thanks to ground-breaking performance improvements in various imaging devices, experts in the medical field can acquire more and better-quality images. At the same time, improvements in network bandwidth and robustness of network devices allow images to be transferred to cloud-based databases in a rapid and timely manner. On the other hand, huge breakthroughs in the semiconductor industry have made storage media cheaper, which has also reduced the cost of storing these images, allowing artificial intelligence algorithms in the medical image field to have more accessible and less expensive to acquire training data. In terms of algorithms, based on a large amount of training data, many neural network-based medical image algorithms such as Convolutional Neural Network, U-net, AlexNet, etc. can help humans to identify diseases with faster speed and higher accuracy by simulating the process of doctors' diagnosis and decision making.

i.1 Brain tumor

For example, the effective use of medical imaging and AI in medicine can be seen in detecting brain tumor. According to Stember & Shalu (2020), Brain cancer can be described as the continuous, abnormal, uncontrolled, self-proliferation of cells in the brain and is one of the most serious cancers. Since the brain is the most important bodily organ of the human body, any small lesion in it has the potential to have a huge impact on physical health, mobility, and thinking ability. Since brain cancer has no obvious symptoms, headache and memory loss are not enough to alert people, so the detection rate of brain cancer is not surprisingly very low. Especially when glioma (Figure 1) appears in the early stage, inexperienced doctors can easily overlook the lesions in medical images. Patients often delay the best time for treatment as a result, allowing their condition to further deteriorate and suffer more pain. Therefore, finding a more detailed test to replace less stable manual ones is a viable way to help patients improve their health status. As artificial intelligence researchers try to apply a wide variety of algorithms to the medical field, the automatic detection of brain tumors by algorithms seems to be an effective solution to this problem.

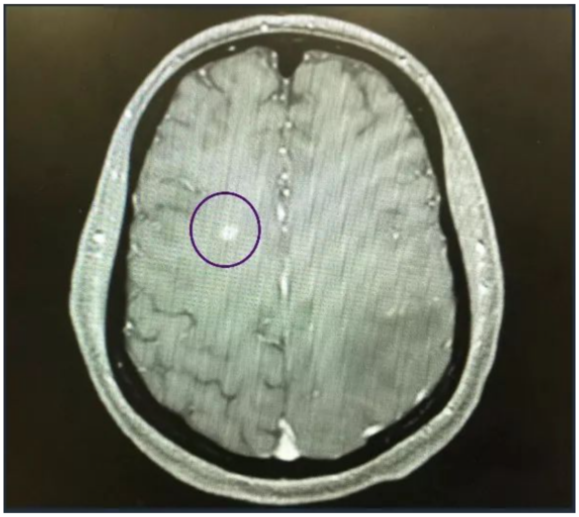

Figure 1: *the shape and size of golioma in a MRI image*

For the field of brain tumor detection, the application of various machine learning and deep learning methods in this field have been rapidly developing since 2015. Shenbagarajan et al. (2016) came up with a method based on Artificial Neural Network. This algorithm uses Active Contour Method to segment the edge of the tumor in images. The Levenberg-Marquardt algorithm is responsible for the classification of the segmented images into three categories: normal, benign tumor, and malignant tumor. As a result of the implementation, this algorithm improved the accuracy by 15 percentage points to 93.74% compared to the SVM-based algorithm commonly used in the past. Sornam et al. (2016) proposed to use K-Means clustering, an unsupervised learning algorithm, to improve the speed of brain tumor image segmentation and recognition. The final algorithm achieved a recognition accuracy of 72%. It is worth mentioning that this algorithm has a 100% detection rate for brain images without tumors on a certain class of images. Ahmmed et al. (2017) tried to use Support Vector Machines to identify tumors from medical images, and later used Artificial Neural Networks to distinguish whether the tumors were benign or malignant, and all four stages of malignant tumors could be identified. This combination allows this algorithm to maintain an accuracy rate of 97.37%.

Stember & Shalu (2020) demonstrated that the artificial intelligence can still perform the task of segmenting brain tumor images even on a very small training set of only ten images. They used unsupervised machine learning methods combined with reinforcement learning to achieve a final recognition rate of 83% on the validation set. Mehrotra et al. (2020) brought the Transfer Learning method into detecting the brain tumor. They used five well-known models of deep learning (AlexNet, ResNet50, ResNet101, GoogleLeNet, and SqueezeNet) to train against medical images related to brain tumors, and later used transfer learning to accelerate the training process to achieve models with similar or higher recognition rates using shorter times. In the end, they trained the model with 99.04% accuracy on this dataset by continuously tuning the hyperparameters at the minute cost of spending 46 minutes and 31 seconds.

Overall, artificial intelligence has gained tremendous progress in detecting brain tumor from medical images since 2015 to date. The enhanced accuracy of models trained by various algorithms on different datasets shows that artificial intelligence is increasingly capable of mimicking doctors to observe, diagnose, and make decisions on brain images. These algorithms not only extract tumor images from the original image with extremely high recognition rates, but also classify malignant tumors into four stages according to their different characteristics. This greatly enhances the speed of medical diagnosis and avoids doctors from making wrong diagnosis due to subjective reasons such as psychological stress and visual bias.

At the same time, the application of some new techniques such as transfer learning and reinforcement learning has made the whole training process increasingly rapid as well as reducing data volume, which speeds up the iteration of medical imaging algorithms and allows artificial intelligence engineers to do more meaningful work in the same effective time. On the other hand, it also promotes wider access to medical services. In some areas where medical resources are scarce, patients or doctors can remotely upload medical images to the cloud and use these algorithms for recognition as a way to compensate for the lack of experienced doctors in certain parts of the world.

i.2 Breast cancer

As one of the most common cancers worldwide, breast cancer causes tremendous suffering for women around the world. 1.67 million new cases of breast cancer were diagnosed worldwide in 2012, which is a quarter of all new cancers that year (Mahersia et al. 2016). Fortunately, the World Health Organization (WHO) pointed out that the five-year survival rate of breast cancer patients can reach 98.8% by timely detection and effective treatment. However, the social stigma specific to women with breast-related conditions, especially in regions such as Asia and Africa, discourages women from going to have their breasts and medical scans of their breasts examined by a male doctor. This largely reduces the chances of being screened for pre-breast cancer. On the other hand, since the pre-cancerous symptoms of breast cancer manifest as microcalcifications (in Figure 2) and lumps, the microcalcification's extremely small size in images makes it difficult to capture this feature with the naked eye, while how to distinguish benign from malignant also troubles medical experts. The lumps vary hugely in size and shape, causing great distress to the process of distinguishing tumors from soft tissues (Mehdy et al. 2017). Therefore, automated detection and screening of medical images of the breast by a machine may be a viable solution in terms of reducing the shame of breast examination in women as well as improving the breast cancer survival rate.

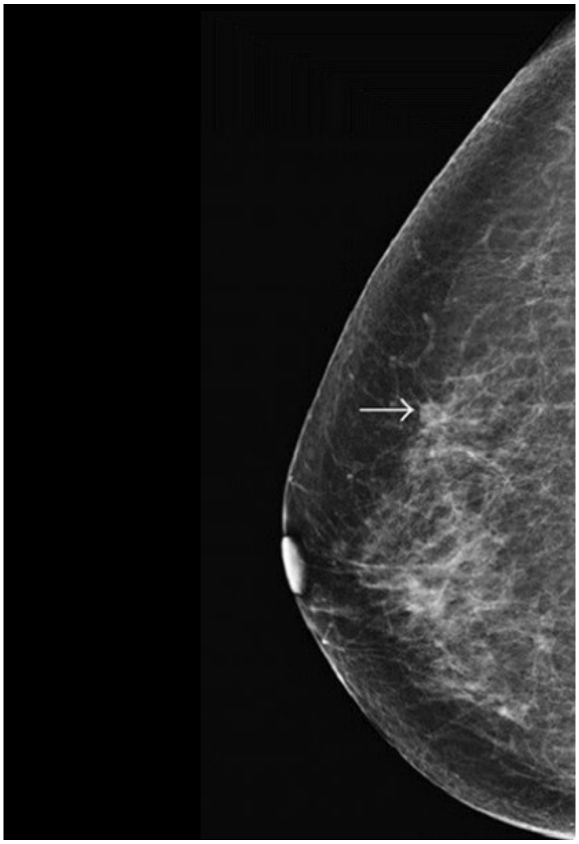

Figure 2: *the microcalcification in a mammogram*

With regards to AI techniques used in detecting breast cancers, Saini & Vijay (2015) applied Feed-Forward Back Propagation and Cascade-Forward Back Propagation Artificial Neural Network to analyse the mammogram images. They used Gray Level Co-occurrence Matrix (GLCM) to extract features from these medical images for distinguishing lumps. Then they evaluate the performance of algorithms by using Mean Square Error (MSE). According to Mahersia et al. (2016), computer-aided diagnosis of medical images of the breast can increase the detection rate of breast cancer by 10 percent, and the use of computers is a more stable and objective way to detect breast cancer than radiologists who are more likely to be influenced by experience as well as subjective judgments.

Mehdy et al. (2017) came up with a method which can enhance and segment mammogram images to find the malignant lumps in the medical image of women's breast. Among other things, their newly discovered statistical features can better identify malignant masses from a variety of different shapes and sizes. According to Zhang et al. (2019), Deep Polynomial Network (DPN) can also be used into diagnosing breast cancer. The special feature is that the dataset used by them is dual-mode, in both shear-wave elastography (SWE) and B-mode ultrasound. The authors evaluated the model performance from four perspectives: sensitivity, specificity, accuracy, and AUC. The model showed high percentages in all four aspects. Romeo et al. (2021) has in recent times validated a test augmented by artificial intelligence, a method that can accurately diagnose breast cancer. They developed 8 radiomics models, combining those models with machine learning. This algorithm achieved the highest accuracy for breast cancer diagnosis (The area of the graph under the variability index curve is 0.983). This method has been evaluated as a promising alternative to physicians for the identification and diagnosis of breast cancer.

These increasingly accurate models, and the application of additional new technologies make identifying medical images of breast cancer increasingly easy for artificial intelligence. Artificial intelligence algorithms can help doctors to identify microcalcifications that are difficult to detect or judge with the naked eye, further improving the recognition accuracy. The more advanced evaluation metrics for the models can also describe the performance, strengths and weaknesses of the algorithms from different perspectives. At the same time, as more medical image data is collected, more and more statistical features are extracted by Artificial Intelligence algorithms and algorithm engineers, and these features can be easily transferred to other similar fields by transferring AI learning methods. In addition, breast cancer is mostly found in women, and because of the prejudice against women in Asian and African societies they can be ashamed of breast-like diseases, especially when doctors of opposite sex are involved in the detection. Fortunately, artificial intelligence algorithms rarely require human involvement in image interpretation, which facilitates early detection and treatment for these women, putting aside their psychological baggage, and promotes gender equality in healthcare.

ii. Face Recognition

Humans are born with the ability to recognize faces, but for computers, it only can distinguish between 0 and 1. How to give this ability to recognize faces to computers is something that many computer scientists have wondered for many years. The face is an important and most significant individual biometric feature, and because it is intuitive and rapid for people to obtain face features, it has an important role in such scenarios as work attendance inspection, security of important institutions, and establishing one's personal identity. But the cost of human resources is rising year by year, and it is a significant expense for employers to set up a department specifically to accomplish these tasks. Furthermore, entrusting these tasks to individuals always renders the execution of the tasks to be subjective. They may mark absent colleagues as attendance due to personal friendships, and allow criminals go into meeting places due to negligence. Therefore, relying only on human beings to perform face recognition in these scenarios carries inaccuracies, unpredictability, and subjectivity.

Thus, finding an intelligent human agent with recognition capabilities and a balance of accuracy, robustness and low complexity becomes tremendously difficult. However, the emergence of artificial intelligence seems to give an answer to this problem. According to the definition of artificial intelligence Russell & Norving (2014), it can think and act rationally like humans.

Artificial intelligence can continuously learn features from the data and experience provided to it by humans and finally behave according to these features. In the field of face recognition, this manifests itself as the AI engineer continuously input images containing faces and labelling the locations of the parts of faces (which may sometimes include the locations of individual feature points on the face). After repeated input, the intelligent agent, i.e., computer, can eventually label the locations of faces in new images with faces.

Zhang et al. (2016) proposed an Adaptive Convolutional Neural Network (ACNN). The greater the network depth with multiple layers results in the greater the performance of the network. However, as more and more networks with many layers are designed, the computer's computational power is no longer able to meet the training requirements of the network, which means that when training a convolutional neural network if a high-precision model is required, artificial intelligence engineers have to adjust hyperparameters and wait for the training to finish.

However, the adaptive convolutional neural network solves this problem to an extent. Adaptive convolutional neural networks can automatically build simple initial network structures with some parameters provided empirically. This greatly reduces the time required for training and strikes a balance between the depth of the network and the training time. As a result, Zhang et al.. trained a model with a higher correct rate (93.33%) in just a quarter of the training time of an ordinary convolutional neural network with a similar structure of depth. In conjunction to that, face recognition has become more cost effective and easily accessible because of the development of cloud computing. Users only need to access the Internet to use the face recognition service in the cloud, and as the calculation process is all carried out in the cloud, users do not need equipment with strong computing power, which brings convenience. In the past, an important reason why artificial intelligence algorithms, including but not limited to face recognition, could not be widely popularized was that it was limited by the computing ability of equipment. Now, thanks to the increased robustness of the network infrastructure and the enhanced computing power in the cloud, face recognition can be made available to small businesses and individuals at much low costs. Zeng et al. (2018) designed a face recognition system based on cloud computing. They used a frame named “AE-FRS” to reduce network latency. The results show that their "AE-FRS" frame is five times more efficient than the traditional local-based face recognition rate when the data volume is extremely large. In addition, face recognition is not only limited to recognizing faces from images, but some scientists have started to try to analyse the emotional state of a person from the expression of the face image. According to Pal et al. (2021), artificial intelligence can now define the expression of a face in a single picture, discerning if someone is feeling a sentiment of happiness, sadness, fear, surprise, anger, and disgust.

Overall, face recognition has gained considerable success in recent years. On the one hand, algorithm engineers can design and train more complex models based on graphics cards with high computing power. On the other hand, they can balance model complexity and training time by reducing the parameters in the model with techniques such as adaptive convolutional neural networks when the computing power cannot meet the demand. In addition, face recognition in a cloud-based system makes it easier to spread, popularize, with lower equipment requirements, wider use of the platform are the fruits of artificial intelligence engineers in this field after deep ploughing.

III. Controversies Artificial Intelligence Brought

It is true that the achievements of artificial intelligence have brought seismic advances to human life and production, but it has also sparked some negative controversies. It is true that the achievements of artificial intelligence have brought disruptive changes to human life and production, but it has also sparked some negative controversies. This section will deal with the ethical and privacy issues involved in the process of medical data acquisition by artificial intelligence, the discrimination issues rendered by artificial intelligence, and the security issues facing deepfake in face recognition, respectively.

i. Data Collection in Medicine

The collection of medical data in the field of artificial intelligence is essentially about using that data to develop algorithms that simulate or assist doctors in the process of diagnosis. Undoubtedly, in recent years, this data has helped predict patient conditions and enhance algorithms for detecting medical images. However, the ethical issues involved in its collection cannot be neglected. According to The Guardian (2019), The “Project Nightingale” transmitted identifiable health information on millions of Americans to Google without the permission of patients or medical data provider. Canales et al. (2020) claims that science without conscience is but the ruin of the soul. They raised questions, whether the ownership of data belongs to the person who generates it or to the person who collects it. If the data is owned by the producer, Google’s actions could even be considered a form of theft.

According to Anom (2020), medical data has already become a kind of goods traded in markets. Individual buyers, research institutions, insurance companies, pharmaceutical companies are willing to pay very high prices for those data such as medical images, patients’ information (including Family history of genetic predisposition, age, gender, etc.), inpatient information and so on. The major question is whether these organizations are doing their part to protect patients’ data. Because of big data, it is easier to describe a person by using user profiling. If these companies do not protect patient data and this data is accessed by advertising companies or hackers, they may face extremely targeted ads, or they may be blackmailed by hackers because of some private diseases.

ii. Consent of medical data provider

According to the definition of consent in criminal law Purdy & Popan (2020), there are three conditions that must be met in order to be considered as the consent of the subject of the act. It based on voluntary agreement; the perpetrator’s mental state is normal; and no fraud and error. For the providers of this medical data, they do not understand what the AI algorithms are doing with their personal data, and even the algorithm engineers themselves do not fully understand what the AI is doing. Therefore, it really is like a black box, the whole process is mysterious and full of unknowns. It is possible that a situation exists where the providers of a genetic disease data only agree to use their data for medical research, and only in an anonymous form, but the artificial intelligence or intelligent agent happens to run through a clustering algorithm that classifies them and their relatives into same categories, and this result happens to be seen by someone who deduces specific individuals by similar family structure, disease type. Does this behaviour violate the patient's level of consent to the extent of data use in a fraudulent manner? This possible error, bias in the process of seeking consent from the data provider led to a misuse of the provider's data and can even be seen as a potential fraud.

Anom (2020) also claimed that, now artificial intelligence is already in hysterectomy equipment. In this scenario, the patient, the device provider, and the physician are not fully aware of exactly what kind of mistakes the intelligent body in the AI might make and how. The patient may have only consented to the removal of site A due to an error in judgment by the AI. However, due to the advice from artificial intelligence the doctor removes site B which is healthy and to which the patient did not consent. In this case, can the device provider, the physician be found to have fraudulently obtained consent from the patient prior to the surgery? If the doctor and the device manufacturer are not responsible, then who should the patient hold accountable? It is clear that the artificial intelligence or intelligent agent is not an independent actor in the legal sense and cannot be held responsible for its actions.

iii. Discrimination in Face Recognition

According to Jackson (2019), discrimination should be defined as unequal treatment of a segment of a group, consciously or unconsciously, for unwarranted reasons. Justifiable reasons, as opposed to improper reasons, include but are not limited to merit, ability, and past performance. Obviously, artificial intelligence itself does not show prejudice or discrimination against any ethnicity, any group. However, the people who create it may introduce prejudice or discrimination in it. According to The New YorkTime (2020), in January 2020 in Michigan, a dark-skinned U.S. citizen who had not violated any laws was arrested for shoplifting. He was handcuffed in front of his home in front of his family, friends, and neighbours. This incident is one of the recent acts of discrimination brought about by artificial intelligence.

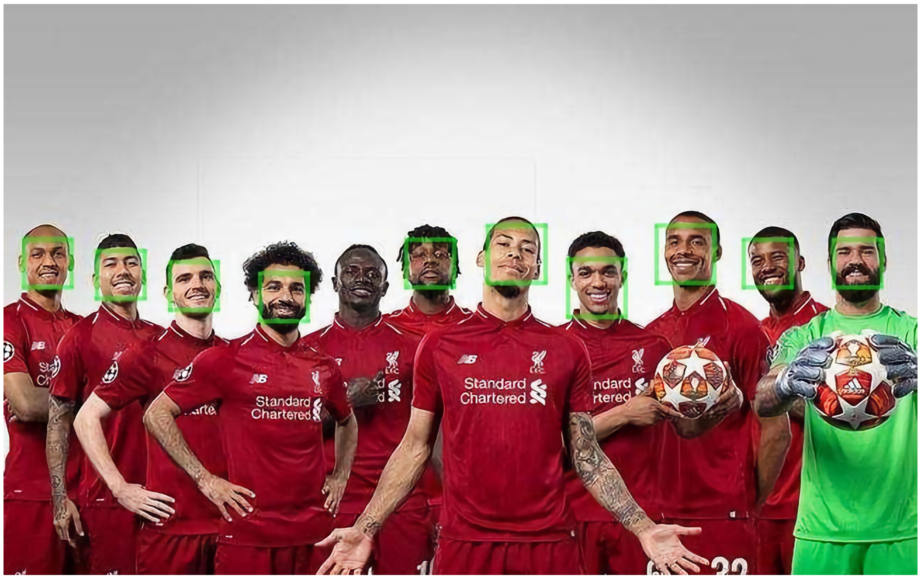

Figure 3: *face recognition algorithm fail to identify African*

The injustice he suffered was actually caused by a mismatch in the facial recognition algorithm. Since most of the current AI research datasets are collected from Caucasians, the data provided by people of other skin colours only account for a small portion of the data, which also causes AI algorithms to tend to match two objects with similar characteristics together or unrecognize minorities (in Figure 3) during the training process because they do not obtain enough features. A Michigan police spokesman said the face recognition algorithms they use in the justice process are so heavily biased that they incorrectly identify black people and Asians, and that this error rates for people of colour is between tens and a hundred times higher than that of Caucasians. In addition, according to Serna et al. (2019), not only skin colour, but also faces with different racial characteristics under the same skin colour, as well as different genders, are treated unfairly by the algorithm. They came up with an algorithm to evaluate the algorithmic discrimination. After testing two popular face recognition algorithms in dataset "DiveFace", they found that both two artificial intelligence algorithms are sensitive to gender, race, and skin colour. As shown in the article, false positives rate is twice as high for women as for men, and up to 1.5 times higher for different ethnic characteristics under the same skin colour.

iv. Deepfake Deceives Face Recognition System

Deepfake is an artificial-intelligence-generated face that does not exist in real life but looks very realistic (in Figure 4). Now it can easily deceive the human visual system and most algorithms in terms of recognizing faces. According to the definition from U.S. Congress (2018), Deepfake is defined as an audio-visual record that is created or modified in such a way that a normal observer would believe it to be a true record of what a person actually said or did, where audio-visual record refers to video material. Due to the existence of open-source community such as GitHub and Gitee, almost everyone can easily access the software, source code of deepfake, so the deepfake, an artificial intelligence application, presents a low threshold.

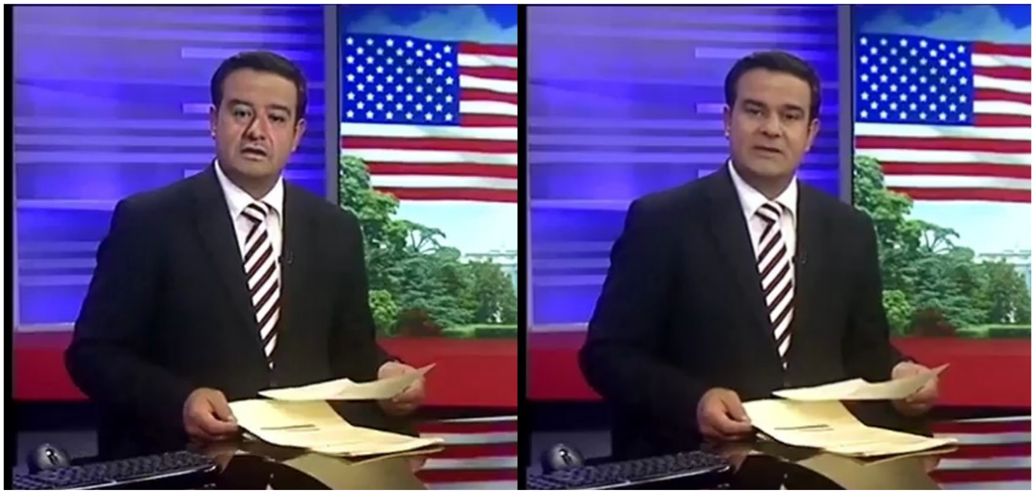

Figure 4: *the image tampered by deepfake (left) and the original image (right)*

Back in 1997, deepfake was born in a video rewriting project, but limited by the computing power of computer hardware, not many people paid attention to deepfake and the problems it could bring at that time. Until the end of 2017, a user on a Reddit forum named Deepfakes replaced the face of a female star of a pornographic film with that of the Wonder Woman super star Gal Gadot. The tampered video by deepfake in YouTube showed that Obama thinks Trump is a "complete dipshit" (BuzzFeedVideo 2018) shocked everyone. Those explosive events brought the controversies about the use of artificial intelligence in this area. Currently, the technology is widely used to produce face-swapping videos of celebrities to create fake news and scandals. In particular, some attackers have replaced the victim's face with the body of a pornographic actor in order to insult, defame and blackmail the victim. On the other hand, this method has also been used as a tool to disseminate political views, posing a huge potential risk to national security in cyberspace, so the review of the authenticity of such video materials is facing a major test.

In addition, Korshunov & Marcel (2019) verified that face recognition algorithms are vulnerable to detect deepfake. They used pre-trained VGG and Facenet as frame of face recognition, both two algorithms achieve SOTA (VGG 98.96, Facenet 99.63). However, the result showed that those two outstanding algorithms cannot effectively distinguish between real faces and faces tampered with by deepfake. This means that intruders can unlock other people's cell phones, pass face recognition in banks, and open face recognition smart door locks at will today when face recognition is widely used in daily life, resulting in reduce of trust between people. It is rather horrible that society is moving towards zero trust situation because of the misuse of technology like deepfake. In an environment where disinformation is rife, people don't have the ability or the will to distinguish the truth of the information they get, and once they believe their trust has been eroded, they will try to be suspicious of any event, which is in fact another kind of interpersonal apathy created by technology. The culprit is actually artificial intelligence. Or is it?

IV. Solution: Convey Morality and Human Conscience to Artificial Intelligence

Gandhi believed that knowledge without character, commerce without morality, science without human conscience was three of the things that would destroy human beings.

In this section, morality and human conscience in AI are jointly seen as cares for human dignity and freedom of human beings in the operation of artificial intelligence applications. Artificial intelligence is essentially just an algorithm, most of the time, the work it does tends to be a projection of AI engineers' intentions through code and machines in real life. Therefore, to give artificial intelligence human conscience, the first step is for the algorithm engineers who create algorithms to act with morality and conscience, and for technology companies to refrain from making unethical development requests of their employees. Based on this, engineers will lead the first step towards transmitting human conscience to technology. However, a more forceful and legally binding solution is necessary as well.

i. Build Consensus

According to the definition of Cambridge Dictionary, "Consensus" should be defined as a rule that everyone in the group abides by. For AI engineers and related technology companies, they need to develop an ethical consensus in the process of designing, creating, and applying artificial intelligence algorithms. This consensus should be accepted and strictly observed by all organizations and individuals involved in artificial intelligence. Considering that AI affects people who use smart devices almost all over the world, such a consensus should be proposed by an international organization. Fortunately, UNESCO has already decided to start drafting a global normative instrument on the ethics of artificial intelligence at the 40th session of the United Nations Educational, Scientific and Cultural Organization (UNESCO) General Conference in November 2019 and released the "Draft text of the Recommendation on the Ethics of Artificial Intelligence" in June 2021.

According to UNESCO (2021), Artificial intelligence practitioners and practicing companies should reach a consensus, artificial intelligence should be developed in a way that does not compromise human dignity, human rights, and fundamental freedoms, and is inclusive of regional differences such as gender, language, religion, and culture, while safeguarding a thriving environment and ecosystem. The draft also focuses on the collection and security of data used by AI, maintaining a balance between collecting data and protecting the privacy of individuals. It is through these universal consensuses that AI practitioners can consciously convey human conscience to AI.

ii. Convey Humanity to Artificial Intelligence

Delivering human conscience to AI means that tech company algorithm engineers must put ethics and the protection of human rights above business and personal selfishness when applying AI. Currently, we can see that some conscientious scientists and engineers are working to use AI to counter that part of AI which is harmful for humans.

Jung et al. (2020) came up with DeepVision algorithm to distinguish deepfake. They found that the flicker pattern of the human eye is significantly different between a normal video and a tampered video, so they used statistical methods and machine learning algorithms to identify the flicker pattern of the eye as a way to distinguish deepfake. Their algorithm can accurately detect seven out of eight different deepfake algorithms with an accuracy rate higher than 87.5%. Kim & Yang (2019) claimed now the autoencoder can anonymize the personal information present and potential in the images. On top of that, these images do not lose the features needed by the AI algorithms, thus achieving a balance between training technological progress and personal privacy protection.

We can see that when AI is given a human conscience and morality by its creators, they perform better on issues such as personal privacy protection and fighting fake news, which are brought about by unregulated AI.

V. Conclusion

Artificial intelligence is just a type of technology, and there is no good or bad in and of itself. The only difference lies in the people who use it. The application of artificial intelligence has proven itself to be competent in many complex and tedious tasks in the field of medical images and face recognition. Although it brings ethical and privacy issues and poses a certain threat to existing face recognition systems, it is also the inherent driving force that promotes the generation and iteration of AI algorithms to solve these problems. At the same time, as the public understands more and more about AI, we have reached or are reaching some consensus about AI, and by algorithm engineers, technology companies will pass these consensuses to AI, conveying human conscience to AI.

References

Ahmmed, R., Swakshar, A. S., Hossain, M. F. & Rafiq, M. A. (2017), Classification of tumors and it stages in brain mri using support vector machine and artificial neural network, *in* '2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)', IEEE, pp. 229–234.

Anom, B. (2020), 'Ethics of big data and artificial intelligence in medicine', *Ethics, Medicine and Public Health* **15**, 100568.

Blasch, E., Pham, T., Chong, C.-Y., Koch, W., Leung, H., Braines, D. & Abdelzaher, T. (2021), 'Machine learning/artificial intelligence for sensor data fusion–opportunities and challenges', *IEEE Aerospace and Electronic Systems Magazine* **36**(7), 80–93.

BuzzFeedVideo (2018), 'You won't believe what obama says in this video!', Website. https://www.youtube.com/watch?v=cQ54GDm1eL0.

Canales, C., Lee, C. & Cannesson, M. (2020), 'Science without conscience is but the ruin of the soul: the ethics of big data and artificial intelligence in perioperative medicine', *Anesthesia and analgesia* **130**(5), 1234.

Cui, Y., Shang, C., Chen, S. & Hao, J. (2019), 'Artificial intelligence overview: The development of ai', *Radio Communications Technology* **45**(3), 225–231.

Deloitte China (2020), 'Scenarios and potentials of al's commercial application in china'. https://www2.deloitte.com/content/dam/Deloitte/cn/Documents/innovation/deloitte-cn-innovation-ai-whitepaper-en-190118.pdf.

Jackson, C. C. (2019), 'Discrimination.', *Salem Press Encyclopedia* .
URL: *https://search.ebscohost.com/login.aspx?direct=trueAuthType=shibdb=ersAN=96397280site=eds-liveauthtype=shibcustid=s3011414*

Jung, T., Kim, S. & Kim, K. (2020), 'Deepvision: Deepfakes detection using human eye blinking pattern.', *IEEE Access, Access, IEEE* **8**, 83144 – 83154.
URL: *https://search.ebscohost.com/login.aspx?direct=trueAuthType=shibdb=edseeeAN=edseee.9072088site=eds-liveauthtype=shibcustid=s3011414*

Kim, T. & Yang, J. (2019), 'Latent-space-level image anonymization with adversarial protector networks', *IEEE Access* **7**, 84992–84999.

Korshunov, P. & Marcel, S. (2019), 'Vulnerability of face recognition to deep morphing', *arXiv preprint arXiv:1910.01933* .

Mahersia, H., Boulehmi, H. & Hamrouni, K. (2016), 'Development of intelligent systems based on bayesian regularization network and neuro-fuzzy models for mass detection in mammograms: A comparative analysis', *Computer methods and programs in biomedicine* **126**, 46–62.

Mehdy, M., Ng, P., Shair, E., Saleh, N. & Gomes, C. (2017), 'Artificial neural networks in image processing for early detection of breast cancer', *Computational and mathematical methods in medicine* **2017**.

Mehrotra, R., Ansari, M., Agrawal, R. & Anand, R. (2020), 'A transfer learning approach for ai-based classification of brain tumors', *Machine Learning with Applications* **2**, 100003.

Pal, S., Mukhopadhyay, S. & Suryadevara, N. (2021), 'Development and progress in sensors and technologies for human emotion recognition', *Sensors* **21**(16), 5554.

Purdy, Elizabeth Rholetter, P. & Popan, Elena, M. (2020), 'Consent (criminal law).', *Salem Press Encyclopedia* .
URL: *https://search.ebscohost.com/login.aspx?direct=trueAuthType=shibdb=ersAN=113931124site=eds-liveauthtype=shibcustid=s3011414*

Razzak, M. I., Naz, S. & Zaib, A. (2018), 'Deep learning for medical image processing: Overview, challenges and the future', *Classification in BioApps* pp. 323–350.

Romeo, V., Clauser, P., Rasul, S., Kapetas, P., Gibbs, P., Baltzer, P., Hacker, M., Woitek, R., Helbich, T. & Pinker, K. (2021), 'Ai-enhanced simultaneous multiparametric 18f-fdg pet/mri for accurate breast cancer diagnosis', *European journal of nuclear medicine and molecular imaging* pp. 1–13.

Russell, S. & Norving, P. (2014), *AI-enhanced simultaneous multiparametric 18F-FDG PET/MRI for accurate breast cancer diagnosis*, Pearson.

Saini, S. & Vijay, R. (2015), Mammogram analysis using feed-forward back propagation and cascade-forward back propagation artificial neural network, *in* '2015 Fifth International Conference on Communication Systems and Network Technologies', pp. 1177–1180.

Serna, I., Morales, A., Fierrez, J., Cebrian, M., Obradovich, N. & Rahwan, I. (2019), 'Algorithmic discrimination: Formulation and exploration in deep learning-based face biometrics', *arXiv preprint arXiv:1912.01842* .

Shenbagarajan, A., Ramalingam, V., Balasubramanian, C. & Palanivel, S. (2016), 'Tumor diagnosis in mri brain image using acm segmentation and ann-lm classification techniques', *Indian Journal of Science and Technology* **9**(1), 1–12.

Sornam, M., Kavitha, M. S. & Shalini, R. (2016), Segmentation and classification of brain tumor using wavelet and zernike based features on mri, *in* '2016 IEEE International Conference on Advances in Computer Applications (ICACA)', IEEE, pp. 166–169.

Stember, J. & Shalu, H. (2020), 'Unsupervised deep clustering and reinforcement learning can accurately segment mri brain tumors with very small training sets', *arXiv preprint arXiv:2012.13321* .

Strümke, I., Slavkovik, M. & Madai, V. (2021), 'The social dilemma in ai development and why we have to solve it', *arXiv preprint arXiv:2107.12977* .

The Guardian (2019), 'Pilkington e google's secret cache of medical data includes names and full details of millions-whistleblower', Website. https://www.theguardian.com/technology/2019/nov/12/google-medical-data-project-nightingale-secret-transfer-us-health-information.

The New YorkTime (2020), 'Wrongfully accused by an algorithm', Website. https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html.

UNESCO (2021), 'Draft text of the recommendation on the ethics of artificial intelligence', Website. https://unesdoc.unesco.org/ark:/48223/pf0000377897.

U.S. Congress (2018), 'Malicious deep fake prohibition act of 2018', Website. https://www.congress.gov/bill/115th-congress/senate-bill/3805/text#idDA40694512F7487981C9819858FA579E.

Zeng, J., Li, C. & Zhang, L.-J. (2018), A face recognition system based on cloud computing and ai edge for iot, *in* 'International Conference on Edge Computing', Springer, pp. 91–98.

Zhang, Q., Song, S., Xiao, Y., Chen, S., Shi, J. & Zheng, H. (2019), 'Dual-mode artificially-intelligent diagnosis of breast tumours in shear-wave elastography and b-mode ultrasound using deep polynomial networks', *Medical engineering & physics* **64**, 1–6.

Zhang, Y., Zhao, D., Sun, J., Zou, G. & Li, W. (2016), 'Adaptive convolutional neural network and its application in face recognition', *Neural Processing Letters* **43**(2), 389–399.