Chapter 1
Data Security Challenges in Deep Neural Network for Healthcare IoT Systems

Edmond S. L. Ho

Abstract With the advancement of IoT technology, more and more healthcare applications were developed in recent years. In addition to the traditional sensor-based systems, image-based healthcare IoT systems become more popular since no specialized sensors are required. Combining with Deep Neural Network (DNN) based automated diagnosis and decision-making systems, it is possible to provide users with 24/7 health monitoring in real life. However, the high computational cost for training DNNs can be a hurdle for developing such kind of powerful systems. While cloud computing can be a feasible solution, uploading training data for the DNN models to the cloud may lead to data security issues. In this chapter, we will review some image-based healthcare IoT systems and discuss some potential risks on data security when training the DNN models on the cloud.

1.1 Introduction

Nowadays, portal devices with high computational power, as well as high-speed data transmission networks [12], are more accessible and affordable to the general public. This certainly enables a wide range of serious applications to be developed to improve our quality of life. In particular, such as environment is suitable for creating healthcare IoT applications which usually consist of capturing health-related data using sensors from the end-user. The data will then be uploaded to the cloud for 1) automated analysis and diagnosis, and/or 2) informing the medical experts or carers in a timely manner if abnormal health conditions are identified.

Examples of healthcare IoT systems in the literature include [17] in which heartbeat sensor, body temperature sensor, room temperature sensor, CO sensor, and CO2 sensor were used for capturing health-related data in a hospital environment to serve

Edmond S. L. Ho
Northumbria University, Sutherland Building, Newcastle upon Tyne, NE1 8ST, United Kingdom, e-mail: e.ho@northumbria.ac.uk

as basic health signs monitoring system for patients. Chatterjee et al. [2] proposed an IoT system for cardiovascular diseases risk assessment of the user based on physiological parameters (age, gender, systolic and diastolic blood pressure, cholesterol, diabetes and smoking habits). Gope and Hwang [11] highlighted the need for a secure healthcare IoT system and proposed BSN-Care which is a secure IoT system based on Body Sensor Network (BSN) technology [22]. In particular, a wide range of sensors was used such a system, including Electrocardiogram (ECG), Electromyography (EMG), Electroencephalography (EEG), Blood Pressure (BP), etc. The secure IoT system is illustrated in Figure 1.1. In addition, both network security (e.g. secure localization, authentication and anonymity) and data security (e.g. data integrity, data freshness and data privacy) are considered when designing the system.

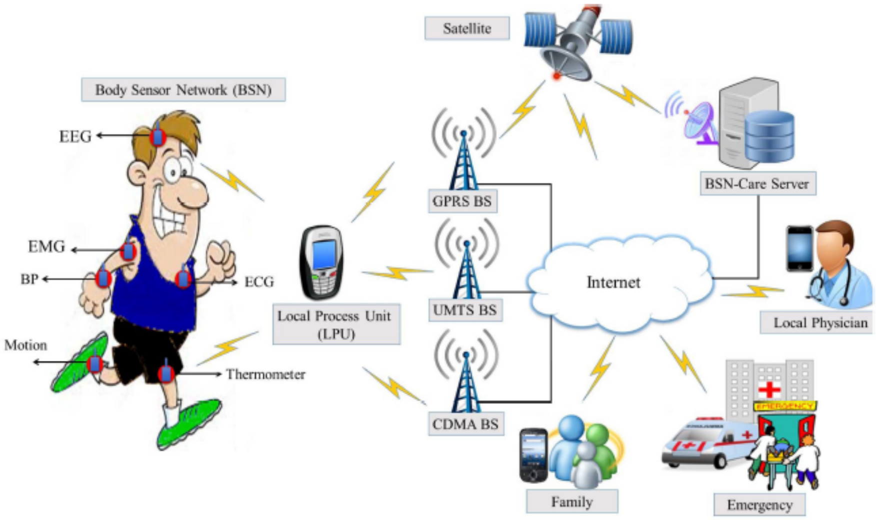


Fig. 1.1: The secure healthcare IoT system using Body Sensor Network (BSN) technology illustrated in [11]. Image reproduced from [11].

Data protection is an important aspect in IoT systems [35] which is not only specific to healthcare applications, but also all other systems that contain and transfer sensitive personal information. For healthcare applications, however, data protection and security become more important due to the fact that highly sensitive personal information will be processed and transferred within the IoT systems. In particular, encryption-based techniques have been widely used for protecting sensitive data to be sent over the network. In [7], Elhoseny et al. proposed a hybrid encryption system for securing secret medical data (such as the diagnostic results). The proposed framework is illustrated in Figure 1.2. Specifically, the encryption method is based on a wide range of methods including Advanced Encryption Standard, Rivest, Shamir, and Adleman algorithms, and the method hides the secret data in a cover image.

In this chapter, we will focus on a less researched area. Firstly, we will review the related research in image-based healthcare IoT systems (Section 1.2) which consists of Deep Neural Network (DNN) as the core technology for automated diagnosis or decision making. Next, different methods for attacking DNNs will be discussed in Section 1.3.

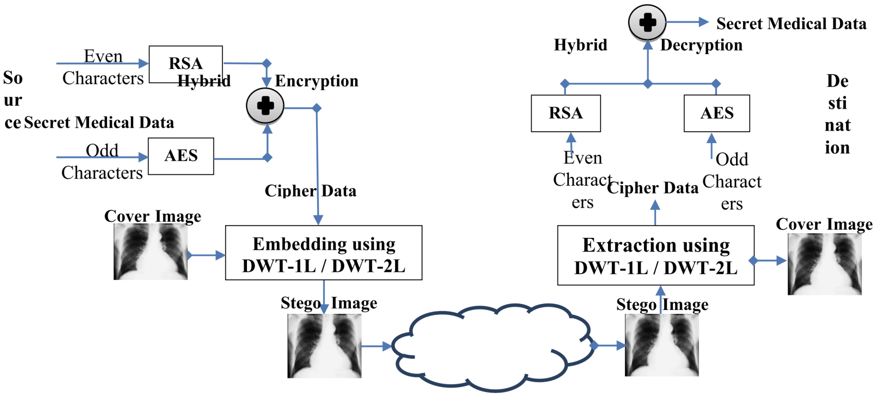


Fig. 1.2: The framework for securing medical data transmission in [7]. Image reproduced from [7].

1.2 Image-based IoT Systems for Automated Diagnosis

Analyzing retinal images provide significant values to the healthcare sector since a wide range of health disorders, such as atherosclerosis, diabetic retinopathy, and congestive heart failure, can be predicted from those images. Recent work by Poplin et al. [28] demonstrated the results on predicting the cardiovascular risk factors from a large retinal fundus photographs dataset, which includes retinal fundus images from 48,101 patients from the UK Biobank (http://www.ukbiobank.ac.uk/about-biobank-uk) and 236,234 patients from EyePACS (http://www.eyepacs.org). The proposed models were validated using images from 12,026 patients from the UK Biobank and 999 patients from EyePACS. In [5], Das R. et al. proposed a low cost and portal healthcare conceptual framework for acquiring retinal fundus images using a Head Mounted Device (HMD). By combining such a framework with a mobile App for Internet connection and deep learning for automated diagnosis, this can potentially be a practical solution for developing countries that have limited access to low-cost retinal image acquisition. In particular, the HMD conceptual image is illustrated in Figure 1.3. The linear actuator will automatically adjust the length to ensure the correct distance between the eyes and lenses can be maintained when capturing the retinal fundus images. Next, the acquired images will be sent to the automated diagnosis system which is essentially a DNN-based image classification framework (such as [28]) as cloud services. Finally, the predicted results will be sent to the medical doctors and/or experts for further analysis. At the time of publishing [5], the system is being translated into a smart healthcare application.

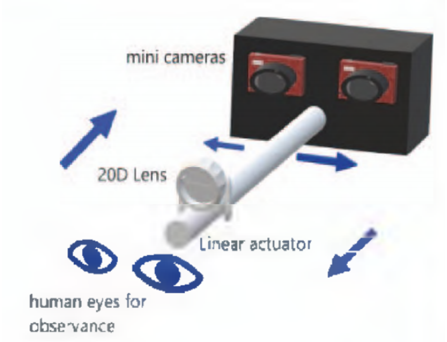


Fig. 1.3: The conceptual illustration of the proposed Head Mounted Device (HMD) for capturing retinal fundus images in [5]. Image reproduced from [5].

Nowadays, the generality of IoT enables a wide range of smart applications to be developed by connecting different types of sensors to the Internet which provide data to be analyzed by automated systems and notify human experts in a timely manner. However, data protection and privacy of the users are crucial factors to ensure the applications are being safe to use. Data encryption is a natural solution for this problem by protecting the data being read by unauthorized parties, but the additional computation cost can greatly degrade the performance of the IoT system especially for systems required to process high volume data. In [19], Jiang et al. proposed an efficient framework for encrypting and diagnosing Diabetic Retinopathy (DR) from retinal images using a camera sensor connected to a Raspberry Pi as a healthcare IoT application. To reduce the computational cost for data encryption, somewhat homomorphic encryption (SHE) is used. Furthermore, parallel homomorphic evaluation is performed by packing multibits into a single ciphertext through single instruction multiple data (SIMD). By this, the computation time for homomorphic evaluation and the transmission time for the cipher text can be reduced simultaneously. For diagnosing DR, density-based clustering [37] is used to classify the retinal images in a highly efficient manner. In summary, the proposed architecture provides a practical solution as a healthcare system with considerations on computational efficiency, privacy protection, storage overhead and communication cost.

Liu et al. [24] proposed combining IoT and artificial intelligence for dental healthcare. In particular, an image-based dental health analysis system is implemented and evaluated. The overview of the proposed framework is illustrated in Figure 1.5. Such an architecture provides users with an in-home dental healthcare platform for sending

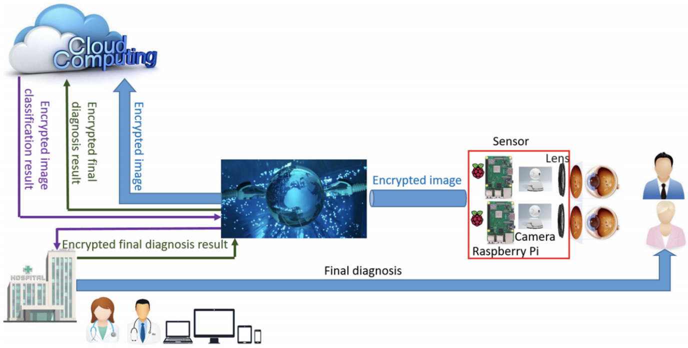

Cloud Computing
Encrypted image classification result
Encrypted final diagnosis result
Encrypted image
Encrypted image
Sensor
Lens
Camera
Raspberry Pi
Encrypted final diagnosis result
HOSPITAL
Final diagnosis

Fig. 1.4: The overview of the proposed framework in [19]. Image reproduced from [19].

color dental images (i.e. RGB images) using a mobile terminal (e.g. smartphone) to the *Smart Dental Services Layer* for detecting dental diseases using a deep learning framework. The *AI Diagnosis of the Teeth* framework is trained using image data annotated with 7 types of dental diseases, including dental caries, dental uorosis, periodontal disease, cracked tooth, dental calculus, dental plaque, and tooth loss. A total of 12,600 clinical images were collected as training samples from 10 private clinics. The core of the AI diagnosis framework is based on Mask R-CNN [14], and image enhancement (such as balancing the dynamic range, edge, and color) is required to achieve better performance. Experimental results show that the diagnosis accuracy ranged from 90.1% to 100% for the 7 types of dental diseases and the mean diagnosis time is reduced by 37.5% in the 1-month trial period in 10 private clinics.

In addition to the aforementioned specialized medical photography IoT systems, there is also a huge potential for typical RGB image (e.g. captured using smartphones [38], camcorders, etc) based IoT systems to be developed as healthcare applications. For example, RGB videos (or RGB image sequences) can be used for predicting cerebral palsy through clinical assessments such as General Movements Assessment (GMA) [6]. The main focus of GMA is assessing the complexity and variability of the general movements of the infant to predict if the nervous system is impaired or not. Most of the current clinical practise still rely on highly trained clinicians to inspect the RGB videos manually. This opens the door for automating this process to:

- reduce subjectivity on the manual assessment
- improve efficiency
- monitor (24/7) the high-risk group such as infants born preterm [13].

A recent work proposed by McCay et al. [25] automate GMA by using computer vision and machine learning techniques. Specifically, skeletal poses are extracted from the video recordings of infant body movements in the pre-processing stage. An example of the skeletal pose estimated from a video frame using OpenPose [1] is il-

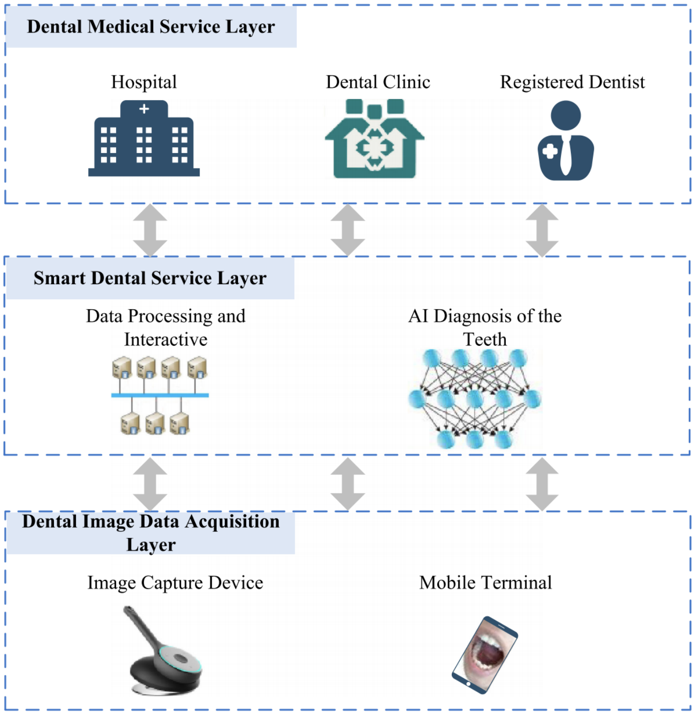

Dental Medical Service Layer
Hospital
Dental Clinic
Registered Dentist
Smart Dental Service Layer
Data Processing and Interactive
AI Diagnosis of the Teeth
Dental Image Data Acquisition Layer
Image Capture Device
Mobile Terminal

Fig. 1.5: The overview of the proposed framework in [24]. Image reproduced from [24].

lustrated in Figure 1.6. Based on the skeletal poses which demonstrated encouraging results in gait analysis [30] as a healthcare application, two new pose-based features, namely Histograms of Joint Orientation 2D (HOJO2D) and Histograms of Joint Displacement 2D (HOJD2D), are proposed to facilitate the classification process to predict if the movement of the infant is considered as *normal* or *abnormal*. In particular, HOJO2D represents the distribution of the orientations of the body parts while HOJD2D represents the distribution of the joint velocity. By fusing such features extracted at each joint, the fused features can be used for representing movement at different levels such as joint-level, limb-level and full body-level. With the use of traditional classifiers such as k-nearest neighbor, Support Vector Machine (SVM) and Ensemble classifier, encouraging result with 91.67% classification accuracy was obtained.

To further improve the classification accuracy, McCay et al. extended the work [26] by proposing 5 DNN architectures as shown in Figure 1.7. Specifically, the new network architectures include fully-connected networks (Figure 1.7a), 1D convolutional neural networks (Figure 1.7b and 1.7c) and 2D convolutional neural networks (Figure 1.7d and 1.7e). Experimental results highlighted that the proposed DNN

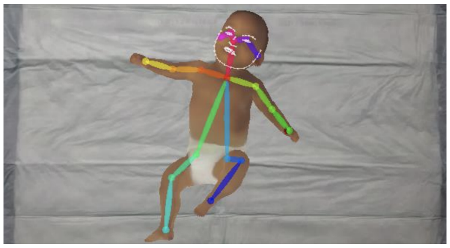
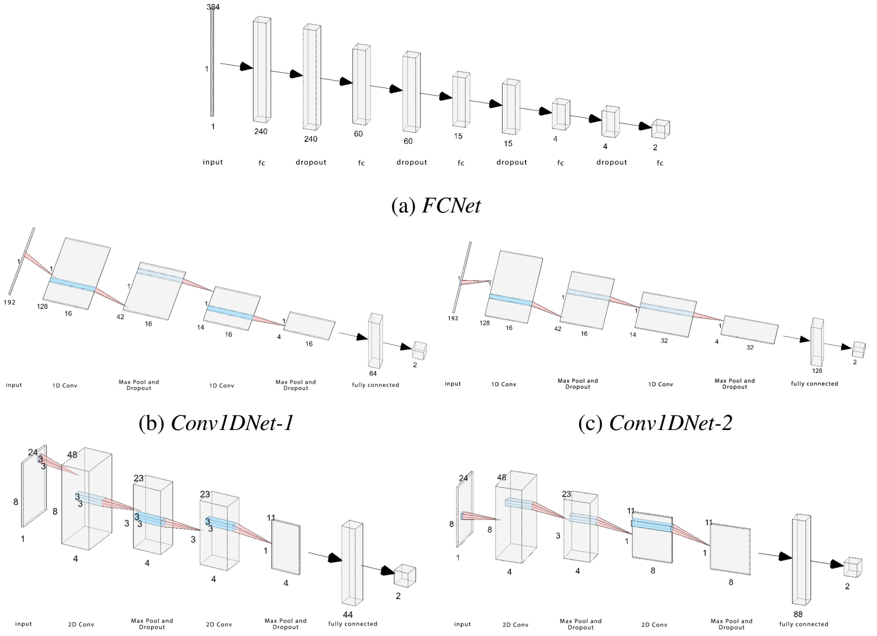


Fig. 1.6: An example of skeletal pose estimation results presented in [25]. Image reproduced from [25].

classifiers are more robust than the traditional classifiers evaluated in [25], and the 1D convolutional neural networks achieved the best performance with the HOJO2D and HOJD2D features.



(a) *FCNet*

(b) *Conv1DNet-1*

(c) *Conv1DNet-2*

(d) *Conv2DNet-1*

(e) *Conv2DNet-2*

Fig. 1.7: The 5 deep neural network architectures for automated prediction of cerebral palsy proposed in [26], including the fully-connected layers based *FCNet*, 1D Convolutional Neural Network based *Conv1DNet-1* and *Conv1DNet-2*, and 2D Convolutional Neural Network based *Conv2DNet-1* and *Conv2DNet-2*. Images reproduced from [26].

.

While McCay et al. [25, 26] mainly focus on the automated diagnostic framework, such an approach can be extended as an IoT application for healthcare by providing the automated diagnosis as cloud services such as the architecture illustrated in Figure 1.9. The use of skeletal pose sequence for the cerebral palsy prediction can certainly lower the risk of leaking sensitive data in case of a data breach since it will be difficult to trace the identity of the subject by only looking at the skeletal poses. To further enhance the accessibility, the video capturing and pose estimation can be implemented as a smartphone App such as the TensorFlow Lite PoseNet (https://www.tensorflow.org/lite/examples, see Figure 1.8). This can possibly be used as in-home monitoring IoT system.

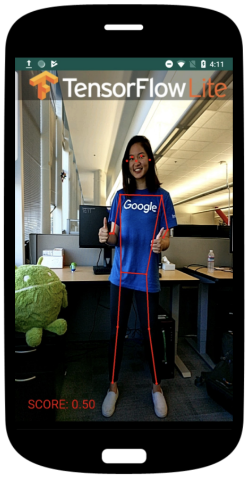


Fig. 1.8: An example of skeletal pose estimation results obtained using TensorFlow Lite PoseNet. Image reproduced from https://www.tensorflow.org/lite/examples.

Pose analysis can also be used in a smart office environment. In [15], Ho et al. proposed an RGB-D camera based monitoring framework (Figure 1.10) to assess the healthiness of the postures of the user in an office environment. Examples of healthy and unhealthy postures are illustrated in Figure 1.11. With the advancement of depth-sensing technology, RGB-D cameras such as Microsoft Kinect become more and more affordable to make it feasible to incorporate such devices in IoT systems. However, the 2.5D data captured from Microsoft Kinect can be noisy and result in incorrect pose extraction (see Figure 1.12). To tackle this problem, Ho et al. [15] proposed to compute an extended set of *reliability values* [31] of the detected joint locations from Microsoft Kinect and incorporate those values into the classification framework. By this, the reliability values will determine the importance of each detected joint locations when classifying the input postures into

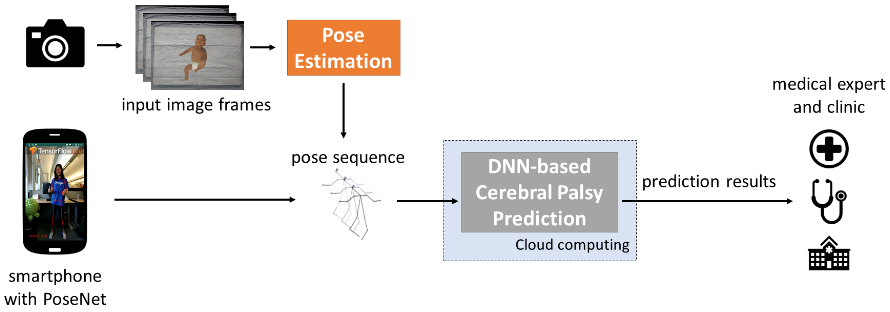


Fig. 1.9: Extending the pose-based cerebral palsy prediction framework to a healthcare IoT system

different healthy/unhealthy classes. Experimental results show that the proposed method outperformed the baselines in the study.

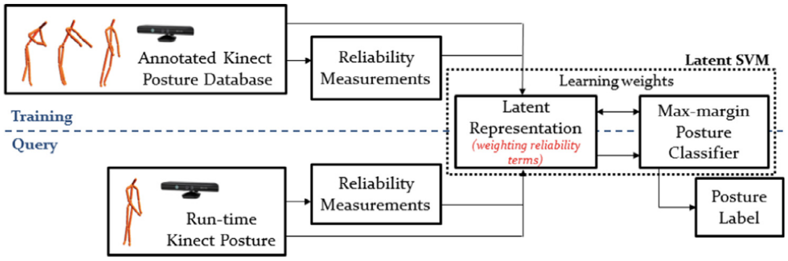
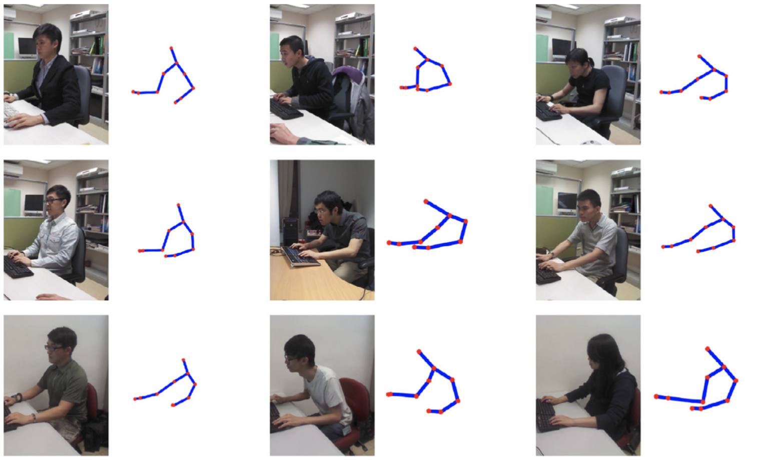


Fig. 1.10: The posture monitoring framework proposed in [15]. Image reproduced from [15].

The aforementioned IoT systems take image or image sequence as input for automated diagnosis. Therefore, the quality of the images is crucial to the success of the systems. However, image quality can be difficult to control when the data is captured at the user-end in which the ideal environment (such as lighting condition, white balance, etc) is not available and the users are inexperienced in taking pictures. To tackle this problem, More et al. [27] proposed a secure Internet of Healthcare Things (IoHT) system which consists of a sparse aware with convolution neural network (SA_CNN) for effective noise removal (Figure 1.13a) and a secure IoHT architecture for medical data storage (Figure 1.13b). The proposed framework was evaluated using various medical modalities and the experimental results show that the new framework outperformed the related work [8, 4, 3, 41] on quantitative measurements such as peak signal to noise ratio (PSNR), structural similarity index (SSIM), and mean squared error (MSE).



(a) Healthy pose

(b) B-2: The back leans forward

(c) B-1: The neck leans forward

Fig. 1.11: Examples of health and unhealthy postures collected in [15]. Image reproduced from [15].

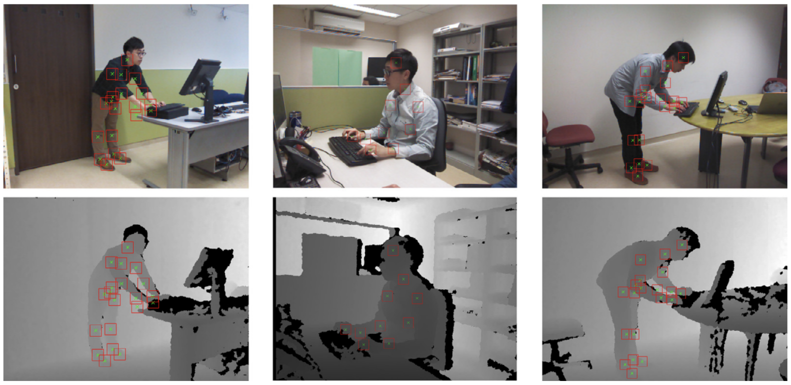
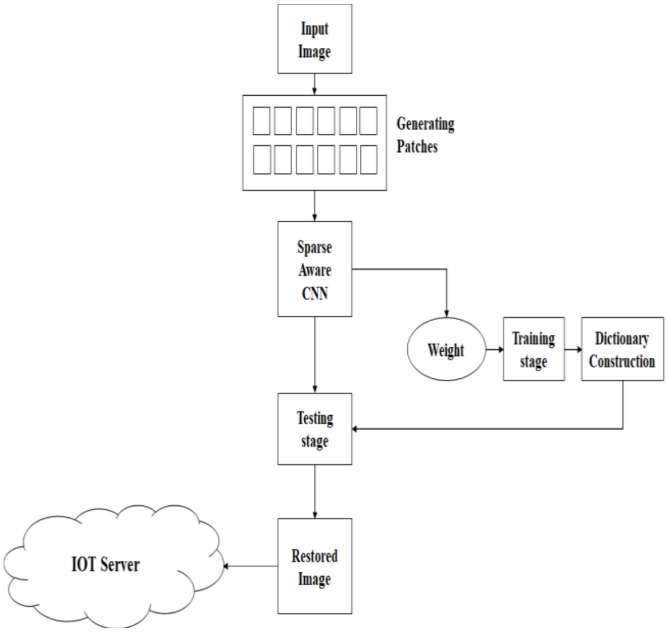
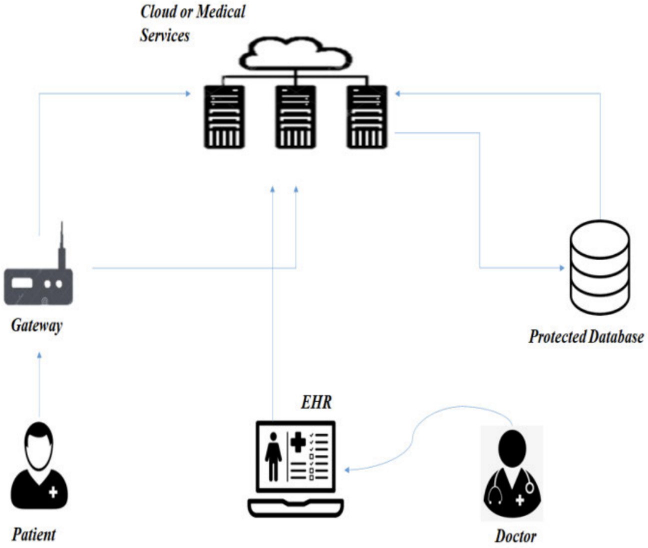


Fig. 1.12: Examples of joint locations detected using the Microsoft Kinect SDK v1. It can be seen that the joint locations (such as the elbow location in the middle column) are detected incorrectly. Image reproduced from [15].



(a) The SA_CNN architecture



(b) The IoHT architecture for medical data storage

Fig. 1.13: The sparse aware convolution Neural Network (SA_CNN) architecture and IoHT architecture for medical data storage proposed in [27]. Images reproduced from [27].

1.3 Data Security Issues in Image-based Deep Learning

With the outstanding performance in using deep learning in different research areas, more and more real-world applications start taking advantage of incorporating DNNs such as the IoT systems discussed in Section 1.2. However, the training process of DNNs usually requires a high volume of training data as well as high computational costs. As a result, cloud service providers including Google, Amazon and Microsoft are offering cloud-based deep learning solutions which can be referred to as Machine Learning as a Service (MLaaS). While such kind of services provides users with the flexibility on the computational resources (i.e. computational power and storage space), uploading training data to remote servers which are managing by external companies may lead to data security problems. Xu et al. [39] recently reviewed some of the data security issues and the solutions available currently, as well as proposed a verifiable and privacy-preserving prediction protocol, namely *SecureNet* for protecting the deep neural networks model and user privacy. In particular, [39] focusing on typical DNN training processing depicted in Figure 1.14 which has an input layer, a number of hidden layers and an output layer. In particular, the output layer contains the prediction (e.g. class label for classification problems). The data security issues are mainly related to attacking the DNN model training process such that wrong prediction will be given as output. In the rest of this section, the terminology from [39] will be used for explaining different types of attacks and solutions.

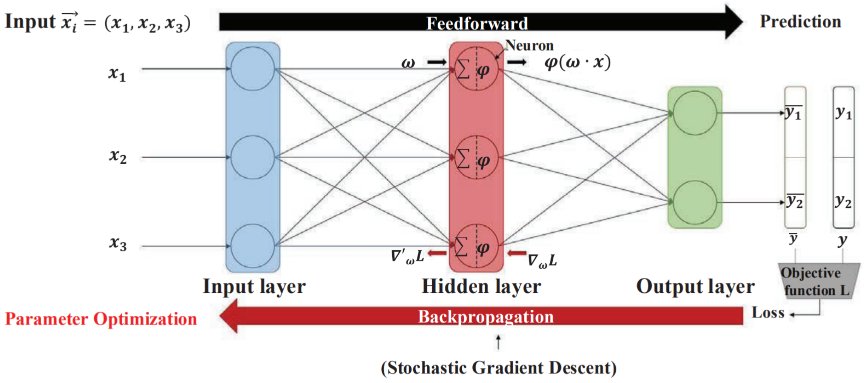


Fig. 1.14: A general Deep neural network training process illustrated in [39]. Image reproduced from [39].

1.3.1 Terminology

1.3.1.1 Centralized Training

This refers to the training process will be done on a single server. Here a single cloud server will be used for getting all training data from the user to complete the DNN training process. In other words, the parameters of the resultant DNN model are computed from the cloud server solely.

1.3.1.2 Collaborative Training

This refers to the training process will be done in a distributed manner. Each user will train the DNN model separately while exchanging the learned model parameters. By this, the ’final’ DNN model is trained in a ’collaborative’ manner.

1.3.1.3 Black-Box Attacker

The attacker can access the DNN model to generate the output (i.e. prediction). However, the attacker does not know the details of the DNN model such as the training data, network architecture and optimization procedures.

1.3.1.4 White-Box Attacker

The attacker can access the DNN model to generate the output (i.e. prediction) as well as the details of the DNN model such as the training data, network architecture and optimization procedures.

1.3.2 Poisoning Attack

The first type of attacks reviewed in [39] is the Poisoning attack. The idea is to have ’poisoned data’ in the training dataset such that that DNN model to be learned will likely generate the wrong prediction. As illustrated in Figure 1.14, the parameters of the DNN model is learned during *backpropagation* in which the gradient is updated in order to minimize the *loss* term(s) of the model. It can be seen that if the training data is being poisoned (or manipulated), the DNN model parameters will not be learned correctly and it may lead to the wrong prediction.

In [18], Jagielski et al. presented a study on poisoning attack against linear regression (Figure 1.15) and proposed an effective defence method, namely *TRIM*, for those attacks. The poisoning attack is done by introducing *poisoning points* into the training data. In particular, both white-box and black-box attackers scenarios were considered and the problems were formulated as a generic bilevel optimization problem. Specifically, for white-box attackers, the outer optimization is responsible for selecting the poisoning points to maximize the loss (opposite to the typical training process which minimizes the loss) while the inner optimization is used for training the regression model parameters based on the poisoned data manipulated by the outer optimization. The black-box attacks follow the same formulation except the training data has to be prepared by the attacker instead of using the 'real' training data.

To defend against the poisoning attacks, the proposed TRIM algorithm [18] not only ignore (or remove) outliers (i.e. likely to be the poisoning points) from the training data as in previous work, but also ignore *inliers* which are the poisoning points that are having a similar data distributions as the real training data. This is achieved by iterative updating the regression model parameters using a subset of training data that has the lowest residuals. Experimental results show that the TRIM algorithm can effectively defence poisoning attacks as the mean square errors (MSEs) are within 1% when compared with the MSEs from the unpoisoned models.

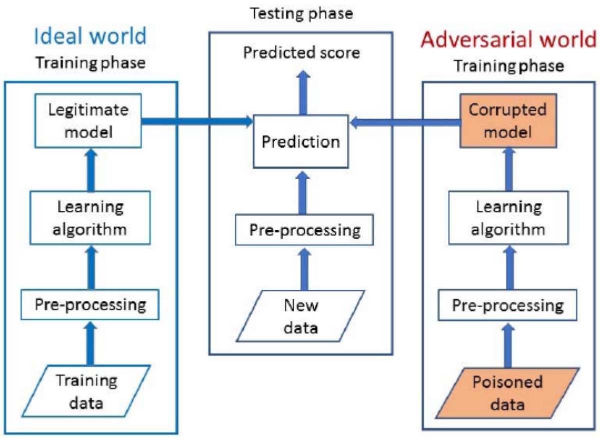


Fig. 1.15: The system architecture for simulating the 'normal' (Ideal world) and 'attack' (Adversarial world) scenarios in [18]. Image reproduced from [18].

Suciu et al. [32] proposed the *FAIL* framework for evaluating the robustness of machine learning models against poisoning and evasion attacks along 4 dimensions: Features, Algorithms, Instances, and Leverage. The framework evaluates a machine learning model by treating different knowledge levels (i.e. availability of model details as in black-box and white-box attacks) as different scenarios and returns the success rates of different attacks. An attack algorithm called *StingRay* is further proposed which demonstrated the effectiveness of bypassing two existing anti-poisoning defences. The main idea behind StingRay is to introduce poisoning points while not affecting the overall classification performance of the DNN model.

Specifically, given a target class of features to be misclassified as the goal of the poisoning attack, StingRay will create poisoning samples that are based on a benign sample from the training data. Next, a subset of the base sample will be replaced by the features of the target class. This process will repeat until the target class of features is being misclassified.

1.3.3 Evasion Attack

The second type of attacks reviewed in [39] is the Evasion attack, in which the goal of the attack is to input carefully crafted samples (i.e. *adversarial examples*) during the prediction (or testing) stage and lead to misclassification. In particular, the adversarial examples can be classified correctly by human but not DNN models. In Figure 1.16, Goodfellow et al. [9] demonstrated an adversarial example which led to misclassification. Specifically, adversarial examples can be generated by adding noise (in [9], the perturbation is based on the gradient of the cost function) to the original input data (e.g. the 'original' panda image in Figure 1.16 left). The resultant adversarial example (Figure 1.16 right) will appear similar to the original image to human eyes, but such input will lead to wrong classification result in DNN models such as classifying this 'panda' image as the 'gibbon' class by GoogLeNet [33].

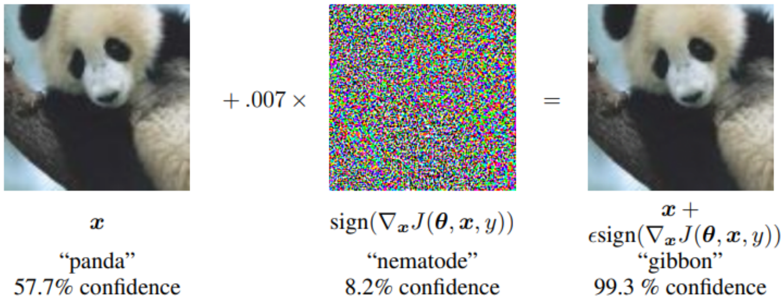
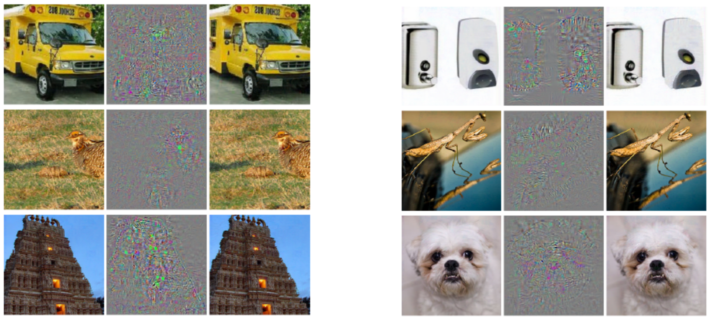

$+ .007 \times$
$=$
$\boldsymbol{x}$
"panda"
57.7% confidence
$\text{sign}(\nabla_{\boldsymbol{x}} J(\boldsymbol{\theta}, \boldsymbol{x}, y))$
"nematode"
8.2% confidence
$\boldsymbol{x} + \epsilon\text{sign}(\nabla_{\boldsymbol{x}} J(\boldsymbol{\theta}, \boldsymbol{x}, y))$
"gibbon"
99.3 % confidence

Fig. 1.16: An adversarial example (right) is created by perturbating the original image (left) based on the gradient of the cost function in [9]. The adversarial example leads to misclassification result on GoogLeNet [33]. Image reproduced from [9].

In the rest of this section, some popular techniques for generating adversarial samples will be discussed. Readers are referred to a recent systematic review [40] as well as the original publications for the details.

Szegedy et al. [34] investigated different properties of neural networks. From their experiments, they found that input-output mappings learned by DNNs are discontinuous. Based on such an interesting property, they proposed an optimization technique to generate adversarial samples. Specifically, the optimization search for the required perturbation by minimizing the probability of images being classified as the correct class. Adversarial samples were generated successfully for attacking various networks (MNIST, QuocNet [23], AlexNet [20]) in the study. Adversarial examples generated for AlexNet [20] are shown in Figure 1.17. It can be seen that the original images and adversarial examples are indistinguishable by a human, but DNN models will misclassify the adversarial examples.



(a) (b)

Fig. 1.17: An adversarial examples generated for AlexNet [20] using the method proposed in [34]. In each set of samples (a and b), the images on the left column are the original images while the right column contains the adversarial examples which lead to misclassification. The middle column illustrates the differences between the original image and the adversarial example magnified by 10 times. Image reproduced from [34].

Goodfellow et al. [9] proposed a method called fast gradient sign method (FGSM) which is more computationally efficient than the expensive optimization used in [34]. Specifically, an adversarial example x' can be generated adding perturbation η to the input image x based on the sign of the gradient of the cost function in the training process:

$$\eta = \epsilon sign(\nabla_x J(\theta, x, y)) \tag{1.1}$$

where y is the target associated with x, θ is the model parameters, $J(\theta, x, y)$ is the cost of training the network and ϵ is used for controlling the strength of the perturbation. An example of the adversarial example generated by this method is visualized in Figure 1.16. Since then, there are variants of the FGSM method proposed in the literature. For example, Tramèr et al. [36] introduced randomness when perturbing the input data, Rozsa et al. [29] replaced the sign of the gradient with the raw gradient value, and Kurakin et al. [21] further extend FGSM to maximizing the target class probability while generating adversarial examples.

1.4 Conclusion

In this chapter, we reviewed a wide range of IoT-based healthcare systems. With the advancement of technology, such as more affordable and portal cameras and better performance smartphones, it is now feasible to capture images from the user-end using portable devices in healthcare IoTs. Combining with image-based deep learning frameworks on the cloud for diagnosis, fully or semi-automated IoT solutions can be realized. While this opens the door for a wide variety of image-based healthcare IoTs to be utilized in real life, we also review some of the potential risks for the image-based deep learning framework which may lead to making the wrong decision. It is advised that IoT solution developers should understand more about the risks and counter those attacks accordingly. For example, Goodfellow et al. [10] and Huang et al. [16] improve the robustness of the DNN models by adding a small number of adversarial examples to the training set iterative. By this, the model to be trained will be more robust again the adversarial examples that may be injected by attackers during the prediction stage.

Acknowledgements This project is supported by the Royal Society (Ref: IES/R1/191147).

References

1. Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. Openpose: Realtime multi-person 2d pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.
2. P. Chatterjee, L. J. Cymberknop, and R. L. Armentano. Iot-based decision support system for intelligent healthcare — applied to cardiovascular diseases. In *2017 7th International Conference on Communication Systems and Network Technologies (CSNT)*, pages 362–366, 2017.
3. Y. Chen and T. Pock. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1256–1272, 2017.
4. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, 2007.
5. P. Das R., G. Rakshitha, I. Juvanna, and D. Venkat Subramanian. Retinal based automated healthcare framework via deep learning. In *2018 Second International Conference on Green Computing and Internet of Things (ICGCIoT)*, pages 93–97, 2018.
6. C. Einspieler and H. F. R. Prechtl. Prechtl's assessment of general movements: A diagnostic tool for the functional assessment of the young nervous system. *Mental Retardation and Developmental Disabilities Research Reviews*, 11(1):61–67, 2005.
7. M. Elhoseny, G. Ramírez-González, O. M. Abu-Elnasr, S. A. Shawkat, N. Arunkumar, and A. Farouk. Secure medical data transmission model for iot-based healthcare systems. *IEEE Access*, 6:20596–20608, 2018.
8. S. Gai and Z. Bao. New image denoising algorithm via improved deep convolutional neural network with perceptive loss. *Expert Systems with Applications*, 138:112815, 2019.
9. I. Goodfellow, J. Shlens, and C. Szegedy. Explaining and harnessing adversarial examples. In *International Conference on Learning Representations*, 2015.

10. I. J. Goodfellow, J. Shlens, and C. Szegedy. Explaining and harnessing adversarial examples, 2015.
11. P. Gope and T. Hwang. Bsn-care: A secure iot-based modern healthcare system using body sensor network. *IEEE Sensors Journal*, 16(5):1368–1376, 2016.
12. B. Gupta and M. Quamara. An overview of internet of things (iot): Architectural aspects, challenges, and protocols. *Concurrency and Computation: Practice and Experience*, 32(21):e4946, 2020. e4946 CPE-18-0159.R1.
13. M. Hafström, K. Källén, F. Serenius, K. Maršál, E. Rehn, H. Drake, U. Ådén, A. Farooqi, K. Thorngren-Jerneck, and B. Strömberg. Cerebral palsy in extremely preterm infants. *Pediatrics*, 141(1), 2018.
14. K. He, G. Gkioxari, P. Dollár, and R. B. Girshick. Mask R-CNN. *CoRR*, abs/1703.06870, 2017.
15. E. S. Ho, J. C. Chan, D. C. Chan, H. P. Shum, Y. ming Cheung, and P. C. Yuen. Improving posture classification accuracy for depth sensor-based human activity monitoring in smart environments. *Computer Vision and Image Understanding*, pages –, 2016.
16. R. Huang, B. Xu, D. Schuurmans, and C. Szepesvari. Learning with a strong adversary, 2016.
17. M. Islam, A. Rahaman, and R. Islam. Development of smart healthcare monitoring system in iot environment. *SN Computer Science*, 1:185, 05 2020.
18. M. Jagielski, A. Oprea, B. Biggio, C. Liu, C. Nita-Rotaru, and B. Li. Manipulating machine learning: Poisoning attacks and countermeasures for regression learning. In *2018 IEEE Symposium on Security and Privacy (SP)*, pages 19–35, 2018.
19. L. Jiang, L. Chen, T. Giannetsos, B. Luo, K. Liang, and J. Han. Toward practical privacy-preserving processing over encrypted data in iot: An assistive healthcare use case. *IEEE Internet of Things Journal*, 6(6):10177–10190, 2019.
20. A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25, pages 1097–1105. Curran Associates, Inc., 2012.
21. A. Kurakin, I. J. Goodfellow, and S. Bengio. Adversarial machine learning at scale. In *International Conference on Learning Representations*, 2017.
22. X. Lai, Q. Liu, X. Wei, W. Wang, G. Zhou, and G. Han. A survey of body sensor networks. *Sensors*, 13(5):5406–5447, Apr 2013.
23. Q. V. Le, M. Ranzato, R. Monga, M. Devin, K. Chen, G. S. Corrado, J. Dean, and A. Y. Ng. Building high-level features using large scale unsupervised learning. In *Proceedings of the 29th International Coference on International Conference on Machine Learning*, ICML'12, page 507–514, Madison, WI, USA, 2012. Omnipress.
24. L. Liu, J. Xu, Y. Huan, Z. Zou, S. C. Yeh, and L. R. Zheng. A smart dental health-iot platform based on intelligent hardware, deep learning, and mobile terminal. *IEEE Journal of Biomedical and Health Informatics*, 24(3):898–906, 2020.
25. K. D. McCay, E. S. L. Ho, C. Marcroft, and N. D. Embleton. Establishing pose based features using histograms for the detection of abnormal infant movements. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5469–5472, July 2019.
26. K. D. McCay, E. S. L. Ho, H. P. H. Shum, G. Fehringer, C. Marcroft, and N. D. Embleton. Abnormal infant movements classification with deep learning on pose-based features. *IEEE Access*, 8:51582–51592, 2020.
27. S. More, J. Singla, S. Verma, Kavita, U. Ghosh, J. J. P. C. Rodrigues, A. S. M. S. Hosen, and I. Ra. Security assured cnn-based model for reconstruction of medical images on the internet of healthcare things. *IEEE Access*, 8:126333–126346, 2020.
28. R. Poplin, A. V. Varadarajan, K. Blumer, Y. Liu, M. McConnell, G. Corrado, L. Peng, and D. Webster. Predicting cardiovascular risk factors in retinal fundus photographs using deep learning. *Nature Biomedical Engineering*, 2018.
29. A. Rozsa, E. M. Rudd, and T. E. Boult. Adversarial diversity and hard positive generation. In *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 410–417, 2016.

30. W. Rueangsirarak, J. Zhang, N. Aslam, E. S. L. Ho, and H. P. H. Shum. Automatic musculoskeletal and neurological disorder diagnosis with relative joint displacement from human gait. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(12):2387–2396, Dec 2018.
31. H. P. H. Shum, E. S. L. Ho, Y. Jiang, and S. Takagi. Real-time posture reconstruction for microsoft kinect. *IEEE Transactions on Cybernetics*, 43(5):1357–1369, 2013.
32. O. Suciu, R. Marginean, Y. Kaya, H. D. III, and T. Dumitras. When does machine learning FAIL? generalized transferability for evasion and poisoning attacks. In *27th USENIX Security Symposium (USENIX Security 18)*, pages 1299–1316, Baltimore, MD, Aug. 2018. USENIX Association.
33. C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–9, 2015.
34. C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus. Intriguing properties of neural networks, 2014.
35. A. Tewari and B. Gupta. Security, privacy and trust of different layers in internet-of-things (iots) framework. *Future Generation Computer Systems*, 108:909–920, 2020.
36. F. Tramèr, A. Kurakin, N. Papernot, I. Goodfellow, D. Boneh, and P. McDaniel. Ensemble adversarial training: Attacks and defenses. In *International Conference on Learning Representations*, 2018.
37. R. Tron, X. Zhou, C. Esteves, and K. Daniilidis. Fast multi-image matching via density-based clustering. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 4077–4086, 2017.
38. W. Wei, E. Ho, K. McCay, R. Damaševičius, R. Maskeliūnas, and A. Esposito. Assessing facial symmetry and attractiveness using augmented reality. *Pattern Analysis and Applications*, 2021.
39. G. Xu, H. Li, H. Ren, K. Yang, and R. H. Deng. Data security issues in deep learning: Attacks, countermeasures, and opportunities. *IEEE Communications Magazine*, 57(11):116–122, 2019.
40. X. Yuan, P. He, Q. Zhu, and X. Li. Adversarial examples: Attacks and defenses for deep learning. *IEEE Transactions on Neural Networks and Learning Systems*, 30(9):2805–2824, 2019.
41. K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017.