

Article

The Application of Design Thinking in Developing a Deep Learning Algorithm for Hip Fracture Detection

Chun-Hsiang Ouyang ¹, Chih-Chi Chen ², Yu-San Tee ¹, Wei-Cheng Lin ³, Ling-Wei Kuo ¹, Chien-An Liao ¹, Chi-Tung Cheng ^{1,*} and Chien-Hung Liao ¹

¹ Department of Trauma and Emergency Surgery, Chang Gung Memorial Hospital, Chang Gung University, Linkou, Taoyuan 33328, Taiwan; detv090@gmail.com (C.-H.O.); b9402011@cgmh.org.tw (Y.-S.T.); m0102@cgmh.org.tw (L.-W.K.); m8407@cgmh.org.tw (C.-A.L.); surgymet@gmail.com (C.-H.L.)

² Department of Rehabilitation and Physical Medicine, Chang Gung Memorial Hospital, Chang Gung University, Linkou, Taoyuan 33328, Taiwan; claudia5477@gmail.com

³ Department of Electrical Engineering, Chang Gung University, Taoyuan 33327, Taiwan; weiclin@mail.cgu.edu.tw

* Correspondence: atong89130@gmail.com; Tel.: +886-3-3281200-3651

Abstract: (1) Background: Design thinking is a problem-solving approach that has been applied in various sectors, including healthcare and medical education. While deep learning (DL) algorithms can assist in clinical practice, integrating them into clinical scenarios can be challenging. This study aimed to use design thinking steps to develop a DL algorithm that accelerates deployment in clinical practice and improves its performance to meet clinical requirements. (2) Methods: We applied the design thinking process to interview clinical doctors and gain insights to develop and modify the DL algorithm to meet clinical scenarios. We also compared the DL performance of the algorithm before and after the integration of design thinking. (3) Results: After empathizing with clinical doctors and defining their needs, we identified the unmet need of five trauma surgeons as “how to reduce the misdiagnosis of femoral fracture by pelvic plain film (PXR) at initial emergency visiting”. We collected 4235 PXRs from our hospital, of which 2146 had a hip fracture (51%) from 2008 to 2016. We developed hip fracture DL detection models based on the Xception convolutional neural network by using these images. By incorporating design thinking, we improved the diagnostic accuracy from 0.91 (0.84–0.96) to 0.95 (0.93–0.97), the sensitivity from 0.97 (0.89–1.00) to 0.97 (0.94–0.99), and the specificity from 0.84 (0.71–0.93) to 0.93(0.90–0.97). (4) Conclusions: In summary, this study demonstrates that design thinking can ensure that DL solutions developed for trauma care are user-centered and meet the needs of patients and healthcare providers.

Keywords: design thinking; artificial intelligence; deep learning; trauma; hip fracture



Citation: Ouyang, C.-H.; Chen, C.-C.; Tee, Y.-S.; Lin, W.-C.; Kuo, L.-W.; Liao, C.-A.; Cheng, C.-T.; Liao, C.-H. The Application of Design Thinking in Developing a Deep Learning Algorithm for Hip Fracture Detection. *Bioengineering* **2023**, *10*, 735. <https://doi.org/10.3390/bioengineering10060735>

Academic Editors: Alan Wang and Andrea Cataldo

Received: 23 April 2023

Revised: 5 June 2023

Accepted: 12 June 2023

Published: 19 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Deep Learning (DL) is a rapidly evolving subcategory of machine learning and is proving to be especially valuable in the healthcare sector [1–4]. DL has revolutionized medical image analysis and has the potential to transform healthcare delivery [5–8]. Previously, medical image analysis relied heavily on manual interpretation by radiologists and physicians, which was often time-consuming and prone to human error. However, DL algorithms have demonstrated remarkable accuracy in analyzing medical images. They have been successful in performing several classification tasks, such as diagnosing lesions, analyzing images, and classifying radiography abnormalities, comparable to or even exceeding those of human experts [9–13].

The integration of DL algorithms in trauma care has the potential to revolutionize the field by fostering the development of innovative solutions [14,15]. Despite its promise, the application of DL algorithms in trauma care is still in its nascent stages when compared to other sectors [16–18]. In trauma management, time is of the essence, and a timely

and accurate diagnosis, coupled with appropriate management, can make a significant difference in the final prognosis. While physician experience is essential for providing high-quality clinical trauma care and treatment [19,20], information from various imaging modalities is also critical [21–23]. The goal of incorporating DL in trauma care is to hasten the diagnostic process, streamline therapeutic decision-making, and ultimately improve patient outcomes [24,25]. Therefore, it is crucial to develop an iterative DL algorithm that can be effectively implemented in clinical practice.

Design thinking is a problem-solving approach that prioritizes empathy, prototyping, and collaboration to generate innovative solutions [26,27]. It is a human-centered approach that places the needs, wants, and preferences of end-users at the center of the design process [28]. Several authors and educators have applied design thinking to medical education [29–32] and medical device development [33,34]. The design thinking process typically involves five steps: empathy, define, ideate, prototype, and test. The empathy stage involves understanding the needs and preferences of the end-users, including clinical physicians and patients. The define stage involves defining the problem and identifying the criteria needed for success. The ideate stage involves brainstorming potential solutions to the problem. The prototype stage involves creating a tangible representation of the solution. Finally, the test stage involves testing the prototype with end-users and gathering feedback to improve the solution.

Although design thinking has shown great breakthroughs in other areas [35–39], there is no previous experience in developing DL algorithms for trauma care. The integration of design thinking into the development of DL algorithms for trauma care has the possibility to modify clinical decision-making and improve patient outcomes. By prioritizing the needs and preferences of end-users, design thinking can lead to the creation of more effective and user-friendly tools [40].

In this study, we aim to demonstrate that incorporating design thinking into the development of DL algorithms for trauma care can help clinical physicians to create the most appropriate algorithm for a clinical scenario. Furthermore, we conducted a multicenter validation to prove the performance of the design thinking-based algorithm.

2. Materials and Methods

We utilized the design thinking process to develop a solution to solve clinical problems. Design thinking can be a useful approach for medical innovation. The process includes five steps: empathy, define, ideate, prototype, and test, with continuous iteration throughout the process [26]. The five steps of design thinking where empathy, define, ideate, prototype, and test (Figure 1).

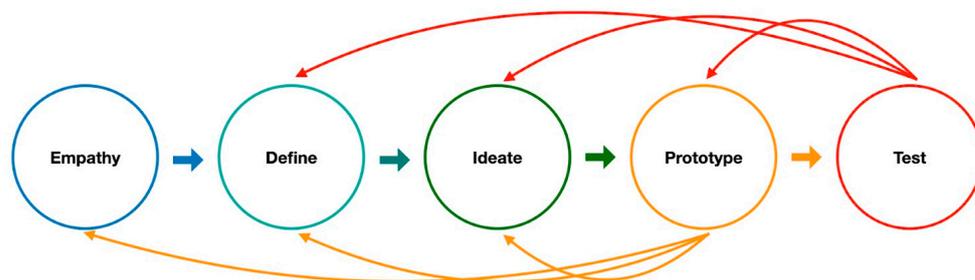


Figure 1. The design thinking process: the five steps include empathy, define, ideate, prototype, and test. Each step can be repeated individually, allowing for an iterative approach to problem-solving.

By utilizing this approach, we were able to develop a solution to address the clinical problems faced by physicians. To ensure that our study was ethical and informed, we interviewed the clinical physicians to understand the clinical unmet need and defined the research question. All of the physician participants were well-informed and consented to participate in the study.

Furthermore, we utilized the Chang Gung Trauma Registry Program (CGTRP) in Chang Gung Memorial Hospital (CGMH), Linkou, Taiwan. Demographic data, medical data, perioperative procedures, hospital procedures, medical imaging findings, follow-up data, and information regarding complications were recorded prospectively in a computerized database. We extracted the data and images of all trauma patients treated from August 2008 to December 2017 at CGMH, which is a level I trauma center for further DL algorithm development. The Internal Review Board of CGMH approved the study with No: 202002343B0.

2.1. Empathize the Clinical Physicians

The first step in design thinking is to empathize with the user. In the context of surgical innovation, this means understanding the needs and challenges of both the patient and the surgical team. We interviewed the physicians before developing the algorithm and labeling the data. The trauma surgeons who work in the trauma bay were interviewed for clinical challenge search and identification. We conducted the interviews to collect any issues about clinical difficulty and to explore the problem.

2.2. Defining the Clinical Unmet Needs and Ideating the Solution for the Clinical Unmet Needs

The research team initiates the process by gathering insights and defining the problem or opportunity. This step will guide the rest of the process of clinical issues and unmet needs into one dominant question that is used to further develop the algorithms. Through brainstorming sessions involving a multidisciplinary team, a broad range of ideas and potential solutions to the problem or opportunity were generated. This involves adjusting the process and concept to integrate the clinical problems with the technological solutions. In the present study, we would like to define the unmet needs and generate ideas for the application of a proper deep learning algorithm to assist clinical physicians in addressing their specific challenges.

2.3. Prototyping the DL Algorithm

The next step is to create a prototype of the potential solution. For this step, we developed a DL algorithm based on the dataset of CGTRP. The PXR was collected from patients registered from May 2008 to December 2016. The demographic and trauma-related data including age, gender, date of injury, mechanism of injury, Abbreviated Injury Scale (AIS), final diagnosis, and outcome, were recorded. We extracted the final diagnosis of hospitalization, operative finding, and the anteroposterior PXR of the patient was acquired from the picture archiving and communication system (PACS) repository.

The architecture of a DL algorithm in this study is based on Xception [41], a convolutional neural network (CNN). The CNN architecture combines depth-wise separable convolutions with residual connections to achieve high accuracy with fewer parameters and reduced computational requirements. The architecture of Xception is based on the idea of using depth-wise separable convolutions, which split the standard convolution operation into two separate operations: depth-wise convolution and pointwise convolution. The depth-wise convolution applies a single convolutional filter to each input channel independently, while the pointwise convolution applies 1×1 convolutions to combine the output channels of the depth-wise convolution. This separation reduces the computational cost of convolutions by factorizing the standard convolution operation. The images were resized into 512×512 pixels on the maximum dimension with zero paddings. The input of the model is the whole image, and the output of the model is a binary classification result representing either hip fracture or no hip fracture. We used 10 image augmentation methods, including blur, brightness, color jitter, contrast adjustment, noise addition, cropping, rotation, shifting, and zooming during the training process (Figure 2).

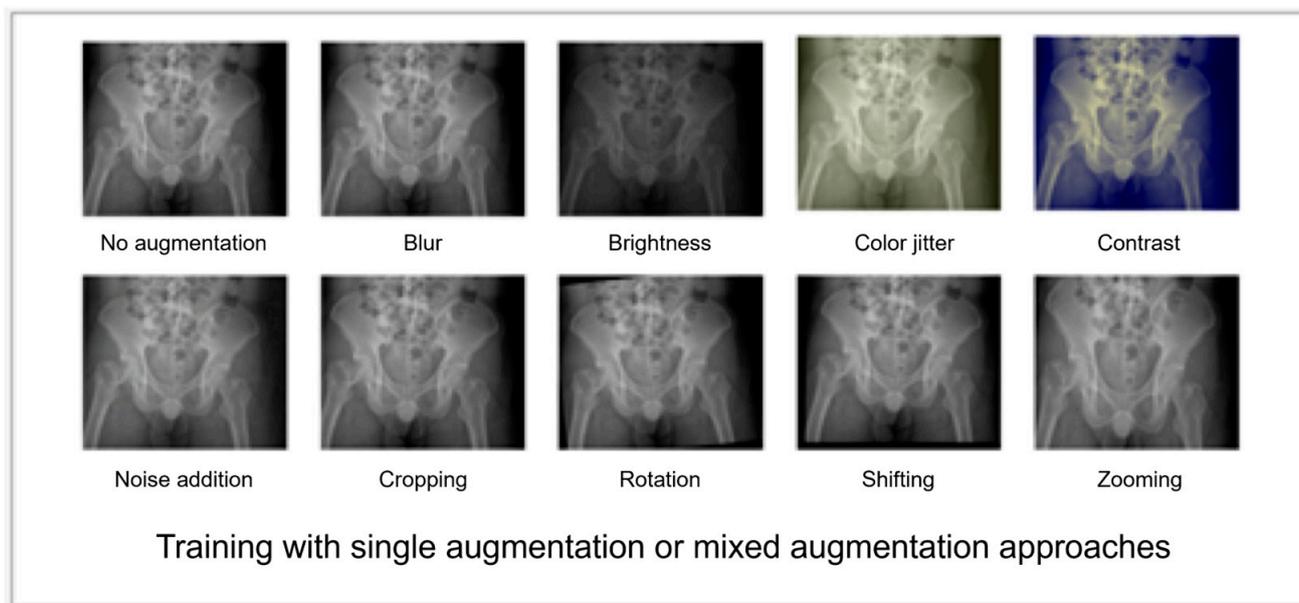


Figure 2. The augmentation methods used for training this deep learning algorithm.

The images after augmentation methods applied were reviewed to check whether the visual features of positive class images can still be recognized. During the training process, the augmented images were randomly added. To test the effect of each image augmentation method, we strategically utilized a diverse set of 10 distinct augmentation methods, each applied individually, to train our model. This approach allowed us to explore the unique impact of each augmentation technique on the performance of the model. Building upon this initial stage, we then proceeded to combine all 10 augmentation methods together, employing a mixed model of training. This subsequent training phase aimed to comprehensively assess and analyze the variations in model performance resulting from the combined use of these diverse augmentation techniques. The models were initialized with the ImageNet pre-trained weight. During the training process, each model was trained with the Adam optimizer using an initial learning rate of 10^{-3} , a batch size of 8, and stopped at 70 epochs. We used Gradient-weighted Class Activation Mapping (Grad-CAM) [42] to visualize the portion of the image that the model focused on to give the prediction. The data of the development dataset from hospital A was divided to 80%, 10%, and 10% as the train, validation, and test dataset, respectively.

We utilized TensorFlow 2.2.0 with Keras 2.3.1, running on Python 3.8.2, to build the structure of the DL algorithm on the Ubuntu 18.04 operating system. The whole training process was run on a GeForce® GTX 1080 Ti GPU. Based on our current hardware configuration, we conducted 70 epochs of training for each model, with each epoch requiring approximately 4 min to complete. Consequently, the entire training process for a single model lasted approximately 280 min.

2.4. Testing, Validating, and Remodeling the Algorithm

The final step is to test the prototype with the testing dataset from CGMH. The test set consisted of 100 PXR including 25 femoral neck fractures, 25 intertrochanteric fractures, and 50 without fracture. We evaluated the performance of the algorithm and incorporated GradCAM visualization to aid user recognition. Following the iterative design-thinking cycle, we carefully reviewed the results and conducted user interviews to gather insights regarding their perception of the algorithm. Specifically, we sought feedback on the performance of the algorithm, the effectiveness of augmentation methods in accounting for variations in clinical environments, and the accuracy of heatmap localization generated by each developed model.

Based on the feedback received, we made enhancements to the algorithm. Instead of solely relying on zoom, horizontal flip, vertical flip, and rotation as augmentation methods in the original version, we introduced multiple additional augmentation techniques to increase the generalizability of the model. Furthermore, we refined the presentation of the algorithm to better cater to user needs and address problem areas identified during the review process.

2.5. Statistical Analysis and Software

The analysis of continuous variables involved using the Kruskal–Wallis rank-sum test for comparison. For categorical variables, the Chi-square test and Fisher’s exact test were used for comparison.

The performance of the model was evaluated with the receiver operating characteristic (ROC) curve. The area under the ROC curve (AUC) was used for evaluation of the performance of the model.

3. Results

3.1. Finding Empathy and Definition

Five trauma surgeons were interviewed to assess their approach to clinical challenges and to identify any unmet needs in handling emergency trauma cases at our hospital. A total of eight unmet needs were identified by the surgeons. These include: (1) detecting a life-threatening hemorrhage that required immediate intervention; (2) predicting critical head injuries that require surgery or intensive care; (3) predicting delayed trauma presentation with secondary complications; (4) determining the need to apply anticoagulant agents and intervention to prevent trauma-induced coagulopathy; (5) detecting subtle fractures in essential x-ray images, such as chest plain films and pelvic plain radiographs (PXR); (6) determining the necessity of surgery in mild blunt or penetrating trauma patients (7) predicting infections or sepsis in patients with multiple trauma; and (8) diagnosing spinal cord injuries that require urgent diagnosis and management. We selected the most frequently reported issue among the surgeons as our primary focus for the study. This issue was detecting subtle fractures in essential x-ray images. Our research team initiated internal discussions and brainstorming sessions, with specific emphasis on the undiagnosed subtle fractures that could result in dismal sequelae such as avascular necrosis, using the principles of design thinking. Finally, the clinical unmet need was defined as “how to reduce the misdiagnosis of femoral fracture by PXR at initial emergency visiting”. To address this unmet need, we developed a visualization-based deep learning algorithm to help clinical physicians identify the presence of possible fractures in PXR in the emergency department.

3.2. Prototyping—Testing Cycle to Improve the Performance of the Algorithms

We collected 4235 anteroposterior PXRs from our hospital, of which 2146 had a hip fracture (51%) from 2008 to 2016. Hip fracture DL detection models based on the Xception CNN were developed using these images. We trained a model without any image augmentation method as a baseline. The developing details were listed in our previous study [43]. In the last layer, we applied Grad-CAM to highlight the lesion predicted by the DL algorithm and explain the process of the algorithm. On the visualization heatmap, all the models highlighted the fractured hip on the PXR image, therefore, we can use the image to assist the clinical doctors to detect the possible area (Figure 3).

After the models developed, the models were directly applied to the emergency department for evaluating the clinical performance of hip fracture detection. The initial results are shown in Table 1. The accuracy, sensitivity, specificity, false-negative rate, and F1 score of the model were 91% ($n = 100$; 95% CI, 84–96%), 97% (95% CI, 89–100%), 84% (95% CI, 71–93%), 2% (95% CI, 0.3–17%), and 0.916 (95% CI, 0.845–0.956), respectively. The details of these aspects were comprehensively described in a previous study [43]. However, unlike previous studies, we implemented this model in a clinical setting and found several limitations in its application. Furthermore, we reevaluated the clinical feedback and

engaged in a process of rethinking to make the necessary modifications to the algorithm to suit its deployment status. Using the feedback from clinical and background calculation, we remodeled the algorithm by using individual image augmentation methods and a model using a mixed method. In addition, the clinical requirement is as follows: “ensure that all the patients discharged from the hospital are free from any fractures”. Therefore, we refine the issue as “increase the sensitivity and reduce the false negative rate”. The cut-off value of the algorithm was adjusted. Furthermore, we refined the presentation from the initial yes/no fracture shift to the visualized presentation, which increased the acceptance of clinical users. Initially, we offered a yes/no fracture presentation to the clinical doctors who did not prefer this suggestion and gave our team their opinion to modify it. After discussion, the algorithm was designed to analyze the X-ray images and provide visual cues to the physician for any possible fracture locations. The physician could then confirm the diagnosis with further investigation or imaging. By these modifications, the performance of the post-design thinking model on the dataset was shown in Table 1. Some of the models trained with augmentation methods, including blur, contrast adjustment, image shifting, and mixed augmentation, performed significantly better than the baseline.

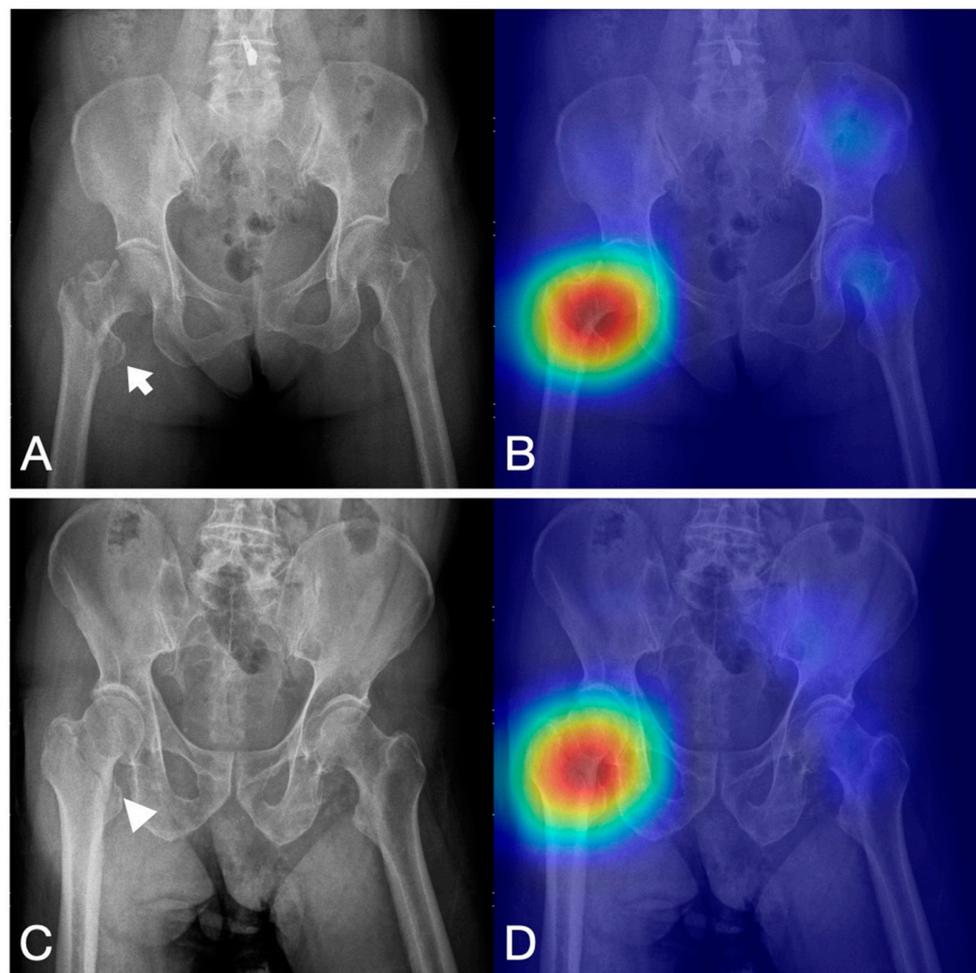


Figure 3. The clinical presentation of the deep learning algorithm for hip fracture detection. (A) The plain pelvic X-ray shows a displaced right intertrochanteric fracture, with the fracture line indicated by an arrow. (B) The heatmap highlights the fracture location, providing clinical doctors with a reference for review. (C) The plain pelvic X-ray shows an occult right intertrochanteric fracture without displacement. The arrowhead indicates the fracture line. (D) The heatmap again indicates the fracture location for review by clinical doctors.

Table 1. The difference in performance before and after the introduction of design thinking for the deep algorithm.

| | Pre DT Model (95%CI) | Post DT Model (95%CI) |
|---------------------|----------------------|------------------------|
| Accuracy | 0.91 (0.84–0.96) | 0.95 (0.93–0.97) |
| Sensitivity | 0.97 (0.89–1.00) | 0.97 (0.94–0.99) |
| Specificity | 0.84 (0.71–0.93) | 0.93 (0.90–0.97) |
| False negative rate | 0.02 (0.003–0.17) | 0.0286 (0.0095–0.0667) |
| F1 score | 0.916 (0.845–0.956) | 0.951 (0.930–0.973) |

The Receiver Operating Characteristic (ROC) curve of the Post-DT model is presented as Figure 4. We can identify the area under the curve as 0.97.

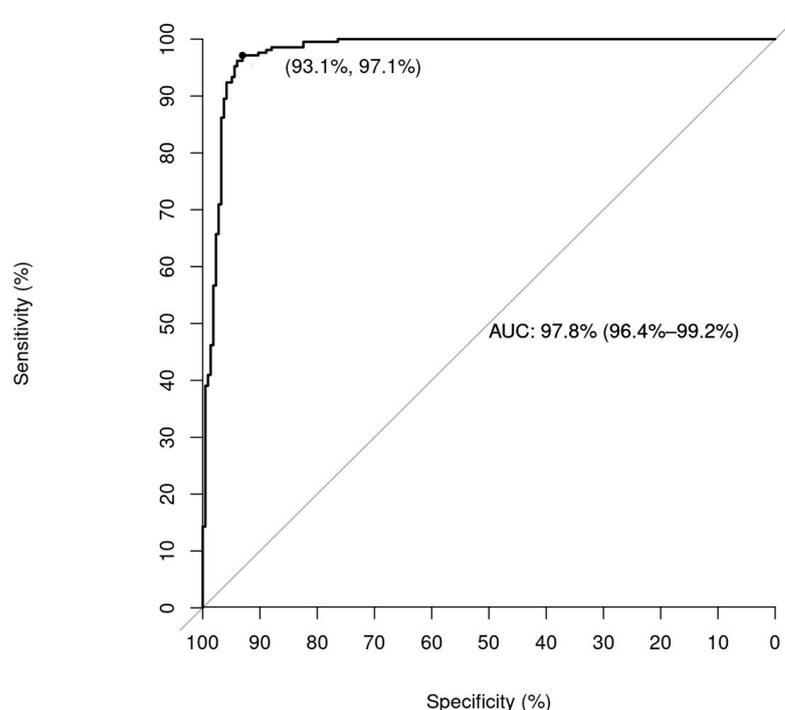


Figure 4. The Receiver Operating Characteristic (ROC) curve of the Post-DT model is presented. The point on the ROC curve signifies the optimal cut-off point based on Youden’s index and also represents the sensitivity and specificity of the model.

4. Discussion

In this study, we presented a project based on Design Thinking to develop DL algorithms that meet the needs of end-users. By identifying their needs quickly and testing the prototypes, we were able to make modifications and deploy the algorithms rapidly. By incorporating feedback from clinical doctors, we were able to improve the diagnostic accuracy from 91% to 95%, the sensitivity from 97% to 97%, and the specificity from 84% to 93%. Trauma care is time-sensitive, and the use of deep learning algorithms can significantly improve the diagnostic rate and time. With the support of these projects, clinicians can make more accurate diagnoses and provide timely treatment to patients.

The management of trauma patients is a complex process, and time is of the essence. Triage plays a crucial role in differentiating patients into appropriate dispositions, and according to advanced trauma life support, it is the first step before the primary survey in managing severely wounded patients. The PXR is a low-cost tool that is essential for detecting possible fractures in trauma patients [44]. With the support of innovative information technology, we have developed AI algorithms capable of accurately classifying

trauma patients within a constrained timeframe. Deep learning technology has been applied to the medical sector for years, and previous works have focused on fracture detection by plain films with excellent performance [6,15,45,46]. Meanwhile, continuous improvement and follow-up results are necessary to ensure optimal performance. In this study, we provide a follow-up and improved performance results of the DL algorithm in fracture detection systems. Once a deep learning algorithm is developed, continuous input of the labeled data can improve its performance [47]. By incorporating increasing data, the algorithm can be trained to recognize new patterns and improve its accuracy. Additionally, clinical feedback is essential for improving and modifying the algorithm to fit clinical needs. This feedback loop is critical for the ongoing development and deployment of deep learning algorithms in clinical practice [48,49].

The continuous improvement of our algorithm is heavily influenced by the insights and expertise of physicians, particularly when it comes to adjusting augmentation methods. Augmentation methods play a crucial role in generating images that serve as valuable training material, enabling us to enhance the generalizability of the algorithm. Physicians provide essential guidance by describing the potential changes that images may undergo in clinical environments, such as blurring and shifting caused by patient movement, as well as variations in brightness and contrast across different machines. These inputs from physicians are invaluable as they help us create augmentation strategies that better reflect real-world scenarios. In the early stages, our algorithm relied solely on traditional augmentation methods, which inevitably limited its generalizability. To address this constraint, we embarked on an iterative journey, systematically integrating various augmentation techniques. Based on the recommendations of the physicians, we constructed model prototypes to experiment with and refine the algorithm. Each iteration allowed us to evaluate the effectiveness of various augmentation methods in enhancing the performance. Through this iterative process, we progressively incorporated a diverse range of augmentation methods. These included techniques for image cropping, rotation, shifting, and zooming, as well as adjustments to blur, brightness, color jitter, contrast adjustment, and noise addition. By integrating multiple augmentation strategies, we aimed to expose the algorithm to a wider range of data variations, thereby enhancing its ability to generalize to unseen scenarios. Once all the augmentation methods were integrated, we observed a significant improvement in performance. The algorithm demonstrated enhanced robustness and adaptability, showcasing its ability to handle various clinical environments and imaging conditions effectively.

The deployment of novel DL algorithms in clinical practice has the potential to revolutionize the field of medicine. However, there are multiple factors that must be considered to ensure successful implementation in real-world settings. Clinical doctors may lack confidence in the new tool, especially if the workings of the algorithm are not explained clearly [50–52]. The heavy computational power and cost of hardware are additional considerations that must be taken into account. Furthermore, most deep learning algorithms only solve one problem separately, which may not be compatible with the clinical workflow and the expectations of frontline physicians [6]. As a result, the deployment of deep learning algorithms in medical scenarios is still in its initial phase. To overcome these challenges, following the design thinking process can help clinical physicians develop deep learning algorithms that are tailored to the specific needs of trauma care [53]. Initially used for business development, the power of design thinking has been recognized for its applicability in various sectors, including the medical field. The Stanford Byers Center for Biodesign has applied design thinking to develop a range of medical innovations [54,55], including surgical tools and devices [56,57]. They have developed a specific process for medical innovation that involves identifying clinical needs, brainstorming potential solutions, creating prototypes, and testing the prototypes with users to refine the design [58]. Some educators use design thinking to arrange and allocate the resources and the coaching system for medical students and young medical residents with great performance [59,60]. In this study, we used design thinking to arrange the process of DL algorithm development.

Involving end-users in the design process is critical to ensure that deep learning algorithms are developed with their needs in mind, leading to increased adoption and usability. By involving frontline physicians in the design process, clinical feedback can be gathered, and the algorithm can be modified to fit clinical needs. This approach can lead to a more effective and efficient deployment of deep learning algorithms in clinical practice. Continuous improvement and follow-up results are necessary to optimize the performance of the algorithms, which is compatible with the core value of design thinking. Rapid iteration of the design thinking cycle and improving the performance of algorithms can enhance the confidence of clinical doctors in the project.

Limitations

While this study demonstrated the efficacy of the design thinking-based deep learning algorithm and its potential for clinical adoption, there were several limitations that should be acknowledged. First, although we attempted to minimize selection bias by randomizing the selection of images, it cannot be completely eliminated due to the nature of the dataset composition. Future studies with larger and more diverse datasets may be required to address this limitation. Second, because the feedback from clinical doctors was an essential part of the study, it was not possible to conduct a double-blind study. This may have introduced bias into the feedback process, potentially affecting the performance of the algorithm. However, we made efforts to minimize this by using a standardized survey form and ensuring that the clinical doctors providing feedback were not involved in the development of the algorithm. Third, as this was a single-center study, the generalizability of the findings may be limited. Further institutional collaborations to expand the data number and improve labeling may be necessary to develop the algorithm according to the global user requirements. This can help ensure that the algorithm is more representative of a broader patient population and healthcare settings, leading to improved performance and more widespread adoption.

5. Conclusions

Design thinking can be used to ensure that the DL solutions developed for trauma care are user-centered and address the needs of patients and healthcare providers. By following the design thinking process, we can gather insights into the needs of users, ideate potential solutions, create prototypes, and test the solutions with users to refine the design. By prioritizing the needs, and preferences of physicians, we can develop DL algorithms that are tailored to the specific needs of trauma care. This can lead to more accurate diagnoses, more efficient treatment, and improved patient outcomes, ultimately contributing to the advancement of trauma care.

Author Contributions: C.-H.O., C.-T.C., and C.-H.L. designed the experiments; C.-T.C., C.-C.C. and Y.-S.T. acquired radiographies for use in the study and provided strategic support; C.-T.C. and W.-C.L. wrote code to achieve different tasks; L.-W.K. and C.-A.L. implemented the annotation tools for data annotation Y.-S.T. and C.-H.L. provided labels for use in measuring algorithm performance; C.-H.O. and W.-C.L. reviewed the computational results; C.-H.O., Y.-S.T., C.-C.C. and C.-T.C. drafted the manuscript; W.-C.L. and C.-H.L. supervised the project. All authors have read and agreed to the published version of the manuscript.

Funding: This publication was supported by the National Science and Technology Council, Taiwan, NTSC- 111-2628-B-182A-010-MY2- (NMRPG3L0321), and Chang Gung Memorial Hospital, grant number CMRPG1K0191.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Chang Gung memorial hospital with no. 202002343B0.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset is not publicly available due to restrictions imposed by the data sharing agreements with the Chang Gung Memorial Hospital Institutional Review Board (IRB). However, a partial dataset can be obtained upon reasonable request to the corresponding authors for academic purposes.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, F.; Casalino, L.P.; Khullar, D. Deep Learning in Medicine—Promise, Progress, and Challenges. *JAMA Intern. Med.* **2019**, *179*, 293–294. [[CrossRef](#)]
2. Ching, T.; Himmelstein, D.S.; Beaulieu-Jones, B.K.; Kalinin, A.A.; Do, B.T.; Way, G.P.; Ferrero, E.; Agapow, P.-M.; Zietz, M.; Hoffman, M.M.; et al. Opportunities and Obstacles for Deep Learning in Biology and Medicine. *J. R. Soc. Interface* **2018**, *15*, 20170387. [[CrossRef](#)]
3. Wang, Z.; Majewicz Fey, A. Deep Learning with Convolutional Neural Network for Objective Skill Evaluation in Robot-Assisted Surgery. *Int. J. Comput. Assist. Radiol. Surg.* **2018**, *13*, 1959–1970. [[CrossRef](#)]
4. Yu, F.; Silva Croso, G.; Kim, T.S.; Song, Z.; Parker, F.; Hager, G.D.; Reiter, A.; Vedula, S.S.; Ali, H.; Sikder, S. Assessment of Automated Identification of Phases in Videos of Cataract Surgery Using Machine Learning and Deep Learning Techniques. *JAMA Netw. Open.* **2019**, *2*, e191860. [[CrossRef](#)]
5. Shen, D.; Wu, G.; Suk, H.-I. Deep Learning in Medical Image Analysis. *Annu. Rev. Biomed. Eng.* **2017**, *19*, 221–248. [[CrossRef](#)]
6. Cheng, C.-T.; Wang, Y.; Chen, H.-W.; Hsiao, P.-M.; Yeh, C.-N.; Hsieh, C.-H.; Miao, S.; Xiao, J.; Liao, C.-H.; Lu, L. A Scalable Physician-Level Deep Learning Algorithm Detects Universal Trauma on Pelvic Radiographs. *Nat. Commun.* **2021**, *12*, 1066. [[CrossRef](#)]
7. Domingues, I.; Pereira, G.; Martins, P.; Duarte, H.; Santos, J.; Abreu, P.H. Using Deep Learning Techniques in Medical Imaging: A Systematic Review of Applications on CT and PET. *Artif. Intell. Rev.* **2020**, *53*, 4093–4160. [[CrossRef](#)]
8. Hinton, G. Deep Learning—A Technology with the Potential to Transform Health Care. *JAMA* **2018**, *320*, 1101–1102. [[CrossRef](#)]
9. Yu, Y.; Xie, Y.; Thamm, T.; Gong, E.; Ouyang, J.; Huang, C.; Christensen, S.; Marks, M.P.; Lansberg, M.G.; Albers, G.W.; et al. Use of Deep Learning to Predict Final Ischemic Stroke Lesions from Initial Magnetic Resonance Imaging. *JAMA Netw. Open.* **2020**, *3*, e200772. [[CrossRef](#)]
10. Lu, M.T.; Ivanov, A.; Mayrhofer, T.; Hosny, A.; Aerts, H.J.W.L.; Hoffmann, U. Deep Learning to Assess Long-Term Mortality from Chest Radiographs. *JAMA Netw. Open.* **2019**, *2*, e197416. [[CrossRef](#)]
11. Cheng, C.-T.; Cai, J.; Teng, W.; Zheng, Y.; Huang, Y.-T.; Wang, Y.-C.; Peng, C.-W.; Tang, Y.; Lee, W.-C.; Yeh, T.-S.; et al. A Flexible Three-Dimensional Heterophase Computed Tomography Hepatocellular Carcinoma Detection Algorithm for Generalizable and Practical Screening. *Hepatol. Commun.* **2022**, *6*, 2901–2913. [[CrossRef](#)] [[PubMed](#)]
12. Jones, R.M.; Sharma, A.; Hotchkiss, R.; Sperling, J.W.; Hamburger, J.; Ledig, C.; O’Toole, R.; Gardner, M.; Venkatesh, S.; Roberts, M.M.; et al. Assessment of a Deep-Learning System for Fracture Detection in Musculoskeletal Radiographs. *NPJ Digit. Med.* **2020**, *3*, 144. [[CrossRef](#)]
13. Olczak, J.; Fahlberg, N.; Maki, A.; Razavian, A.S.; Jilert, A.; Stark, A.; Sköldenberg, O.; Gordon, M. Artificial Intelligence for Analyzing Orthopedic Trauma Radiographs: Deep Learning Algorithms—Are They on Par with Humans for Diagnosing Fractures? *Acta Orthop.* **2017**, *88*, 581–586. [[CrossRef](#)] [[PubMed](#)]
14. Monteiro, M.; Newcombe, V.F.J.; Mathieu, F.; Adatia, K.; Kamnitsas, K.; Ferrante, E.; Das, T.; Whitehouse, D.; Rueckert, D.; Menon, D.K.; et al. Multiclass Semantic Segmentation and Quantification of Traumatic Brain Injury Lesions on Head CT Using Deep Learning: An Algorithm Development and Multicentre Validation Study. *Lancet Digit. Health* **2020**, *2*, e314–e322. [[CrossRef](#)]
15. Gipson, J.; Tang, V.; Seah, J.; Kavvoudias, H.; Zia, A.; Lee, R.; Mitra, B.; Clements, W. Diagnostic Accuracy of a Commercially Available Deep-Learning Algorithm in Supine Chest Radiographs Following Trauma. *Br. J. Radiol.* **2022**, *95*, 20210979. [[CrossRef](#)] [[PubMed](#)]
16. Chiu, I.-M.; Lin, C.-H.R.; Yau, F.-F.F.; Cheng, F.-J.; Pan, H.-Y.; Lin, X.-H.; Cheng, C.-Y. Use of a Deep-Learning Algorithm to Guide Novices in Performing Focused Assessment with Sonography in Trauma. *JAMA Netw. Open.* **2023**, *6*, e235102. [[CrossRef](#)]
17. Regnard, N.-E.; Lanseur, B.; Ventre, J.; Ducarouge, A.; Clovis, L.; Lassalle, L.; Lacave, E.; Grandjean, A.; Lambert, A.; Dallaudière, B.; et al. Assessment of Performances of a Deep Learning Algorithm for the Detection of Limbs and Pelvic Fractures, Dislocations, Focal Bone Lesions, and Elbow Effusions on Trauma X-Rays. *Eur. J. Radiol.* **2022**, *154*, 110447. [[CrossRef](#)]
18. Adebayo, O.; Bhuiyan, Z.A.; Ahmed, Z. Exploring the Effectiveness of Artificial Intelligence, Machine Learning and Deep Learning in Trauma Triage: A Systematic Review and Meta-Analysis. *Lancet* **2022**, preprint. [[CrossRef](#)]
19. Haas, B.; Jurkovich, G.J.; Wang, J.; Rivara, F.P.; Mackenzie, E.J.; Nathens, A.B. Survival Advantage in Trauma Centers: Expeditious Intervention or Experience? *J. Am. Coll. Surg.* **2009**, *208*, 28–36. [[CrossRef](#)]
20. Haut, E.R.; Chang, D.C.; Efron, D.T.; Cornwell, E.E. 3rd Injured Patients Have Lower Mortality When Treated by “Full-Time” Trauma Surgeons vs. Surgeons Who Cover Trauma “Part-Time”. *J. Trauma.* **2006**, *61*, 272–278; discussion 278–279. [[CrossRef](#)]
21. Gay, D.A.T.; Miles, R.M. Use of Imaging in Trauma Decision-Making. *J. R. Army Med. Corps* **2011**, *157*, S289–S292. [[CrossRef](#)]
22. Peytel, E.; Menegaux, F.; Cluzel, P.; Langeron, O.; Coriat, P.; Riou, B. Initial Imaging Assessment of Severe Blunt Trauma. *Intensive Care Med.* **2001**, *27*, 1756–1761. [[CrossRef](#)] [[PubMed](#)]

23. Long, B.; April, M.D.; Summers, S.; Koyfman, A. Whole Body CT versus Selective Radiological Imaging Strategy in Trauma: An Evidence-Based Clinical Review. *Am. J. Emerg. Med.* **2017**, *35*, 1356–1362. [CrossRef] [PubMed]
24. Choi, J.; Alawa, J.; Tennakoon, L.; Forrester, J.D. DeepBackRib: Deep Learning to Understand Factors Associated with Readmissions after Rib Fractures. *J. Trauma. Acute Care Surg.* **2022**, *93*, 757–761. [CrossRef] [PubMed]
25. Rojas-Muñoz, E.; Couperus, K.; Wachs, J.P. The AI-Medic: An Artificial Intelligent Mentor for Trauma Surgery. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2021**, *9*, 313–321. [CrossRef]
26. Rowe, P.G. *Design Thinking*; MIT Press: Cambridge, MA, USA, 1991; ISBN 9780262680677.
27. Brown, T. Design Thinking. *Harv. Bus. Rev.* **2008**, *86*, 84–92, 141.
28. Dorst, K. The Core of “design Thinking” and Its Application. *Des. Stud.* **2011**, *32*, 521–532. [CrossRef]
29. Deitte, L.A.; Omary, R.A. The Power of Design Thinking in Medical Education. *Acad. Radiol.* **2019**, *26*, 1417–1420. [CrossRef]
30. Badwan, B.; Bothara, R.; Latijnhouwers, M.; Smithies, A.; Sandars, J. The Importance of Design Thinking in Medical Education. *Med. Teach.* **2018**, *40*, 425–426. [CrossRef]
31. Sandars, J.; Goh, P.-S. Design Thinking in Medical Education: The Key Features and Practical Application. *J. Med. Educ. Curric. Dev.* **2020**, *7*, 2382120520926518. [CrossRef]
32. van de Grift, T.C.; Kroeze, R. Design Thinking as a Tool for Interdisciplinary Education in Health Care. *Acad. Med.* **2016**, *91*, 1234–1238. [CrossRef] [PubMed]
33. Carroll, N.; Richardson, I. Aligning Healthcare Innovation and Software Requirements through Design Thinking. In Proceedings of the International Workshop on Software Engineering in Healthcare Systems, Austin, TX, USA, 14–22 May 2016; Association for Computing Machinery: New York, NY, USA; pp. 1–7.
34. Zupan, B.; Pakrac, V.; Nabergoj, A.S. Using Design Thinking to Develop a New Product: Case Study of Developing a Novel Medical Device. In *Design Thinking*; CRC Press: Boca Raton, FL, USA, 2022; Available online: <https://www.taylorfrancis.com/chapters/edit/10.1201/9781003189923-12/using-design-thinking-develop-new-product-bla%20C5%BE-zupan-vladimir-pakrac-anja-svetina-nabergoj> (accessed on 18 April 2023).
35. Brooks, F.P., Jr. *The Design of Design: Essays from a Computer Scientist*; Pearson Education: London, UK, 2010; ISBN 9780321702067.
36. Martin, R.; Martin, R.L. *The Design of Business: Why Design Thinking Is the Next Competitive Advantage*; Harvard Business Press: Cambridge, MA, USA, 2009; ISBN 9781422177808.
37. Panke, S. Design Thinking in Education: Perspectives, Opportunities and Challenges. *Open. Educ. Stud.* **2019**, *1*, 281–306. [CrossRef]
38. Mintrom, M.; Luetjens, J. Design Thinking in Policymaking Processes: Opportunities and Challenges. *Aust. J. Publ. Adm.* **2016**, *75*, 391–402. [CrossRef]
39. Watson, A.D. Design Thinking for Life. *Art. Educ.* **2015**, *68*, 12–18. [CrossRef]
40. Scholten, H.; Granic, I. Use of the Principles of Design Thinking to Address Limitations of Digital Mental Health Interventions for Youth: Viewpoint. *J. Med. Internet Res.* **2019**, *21*, e11528. [CrossRef]
41. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–16 July 2017; pp. 1251–1258.
42. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-Cam: Visual Explanations from Deep Networks via Gradient-Based Localization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 618–626.
43. Cheng, C.-T.; Ho, T.-Y.; Lee, T.-Y.; Chang, C.-C.; Chou, C.-C.; Chen, C.-C.; Chung, I.-F.; Liao, C.-H. Application of a Deep Learning Algorithm for Detection and Visualization of Hip Fractures on Plain Pelvic Radiographs. *Eur. Radiol.* **2019**, *29*, 5469–5477. [CrossRef]
44. Shenton, A.; Choudhary, S. The Emergency Radiology of Pelvic Trauma. *Traumatology* **2014**, *16*, 279–291. [CrossRef]
45. Açıcı, K.; Sümer, E.; Beyaz, S. Comparison of Different Machine Learning Approaches to Detect Femoral Neck Fractures in X-Ray Images. *Health Technol.* **2021**, *11*, 643–653. [CrossRef]
46. Olczak, J.; Emilson, F.; Razavian, A.; Antonsson, T.; Stark, A.; Gordon, M. Ankle Fracture Classification Using Deep Learning: Automating Detailed AO Foundation/Orthopedic Trauma Association (AO/OTA) 2018 Malleolar Fracture Identification Reaches a High Degree of Correct Classification. *Acta Orthop.* **2021**, *92*, 102–108. [CrossRef]
47. Pianykh, O.S.; Langs, G.; Dewey, M.; Enzmann, D.R.; Herold, C.J.; Schoenberg, S.O.; Brink, J.A. Continuous Learning AI in Radiology: Implementation Principles and Early Applications. *Radiology* **2020**, *297*, 6–14. [CrossRef]
48. Ngiam, K.Y.; Khor, I.W. Big Data and Machine Learning Algorithms for Health-Care Delivery. *Lancet Oncol.* **2019**, *20*, e262–e273. [CrossRef] [PubMed]
49. Budd, S.; Robinson, E.C.; Kainz, B. A Survey on Active Learning and Human-in-the-Loop Deep Learning for Medical Image Analysis. *Med. Image Anal.* **2021**, *71*, 102062. [CrossRef] [PubMed]
50. Kim, J.H.; Han, S.G.; Cho, A.; Shin, H.J.; Baek, S.-E. Effect of Deep Learning-Based Assistive Technology Use on Chest Radiograph Interpretation by Emergency Department Physicians: A Prospective Interventional Simulation-Based Study. *BMC Med. Inform. Decis. Mak.* **2021**, *21*, 311. [CrossRef]
51. Oh, S.; Kim, J.H.; Choi, S.-W.; Lee, H.J.; Hong, J.; Kwon, S.H. Physician Confidence in Artificial Intelligence: An Online Mobile Survey. *J. Med. Internet Res.* **2019**, *21*, e12422. [CrossRef] [PubMed]

52. Miller, D.D.; Brown, E.W. Artificial Intelligence in Medical Practice: The Question to the Answer? *Am. J. Med.* **2018**, *131*, 129–133. [[CrossRef](#)]
53. Auernhammer, J.; Roth, B. The Origin and Evolution of Stanford University's Design Thinking: From Product Design to Design Thinking in Innovation Management. *J. Prod. Innov. Manag.* **2021**, *38*, 623–644. [[CrossRef](#)]
54. Zenios, S.; Makower, J.; Yock, P.; Denend, L.; Brinton, T.J.; Kumar, U.N.; Krummel, T.M. *Biodesign: The Process of Innovating Medical Technologies*; Cambridge University Press: Cambridge, MA, USA, 2010; ISBN 9780521517423.
55. Wall, J.; Hellman, E.; Denend, L.; Rait, D.; Venook, R.; Lucian, L.; Azagury, D.; Yock, P.G.; Brinton, T.J. The Impact of Postgraduate Health Technology Innovation Training: Outcomes of the Stanford Biodesign Fellowship. *Ann. Biomed. Eng.* **2017**, *45*, 1163–1171. [[CrossRef](#)]
56. Brinton, T.J.; Kurihara, C.Q.; Camarillo, D.B.; Pietzsch, J.B.; Gorodsky, J.; Zenios, S.A.; Doshi, R.; Shen, C.; Kumar, U.N.; Mairal, A.; et al. Outcomes from a Postgraduate Biomedical Technology Innovation Training Program: The First 12 Years of Stanford Biodesign. *Ann. Biomed. Eng.* **2013**, *41*, 1803–1810. [[CrossRef](#)]
57. Wall, J.; Wynne, E.; Krummel, T. Biodesign Process and Culture to Enable Pediatric Medical Technology Innovation. *Semin. Pediatr. Surg.* **2015**, *24*, 102–106. [[CrossRef](#)]
58. Dutta, N.; Dhar, D. From Industrial Design to Healthcare Innovation—A Comparative Study on the Role of User-Centered Design and Stanford Biodesign Process. In *Design for Tomorrow*; Springer: Singapore, 2021; Volume 3, pp. 665–677. [[CrossRef](#)]
59. Gottlieb, M.; Wagner, E.; Wagner, A.; Chan, T. Applying Design Thinking Principles to Curricular Development in Medical Education. *AEM Educ. Train.* **2017**, *1*, 21–26. [[CrossRef](#)]
60. Madson, M.J. Making Sense of Design Thinking: A Primer for Medical Teachers. *Med. Teach.* **2021**, *43*, 1115–1121. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.