International Journal of
Applied Data Science in Engineering and Health
https://ijadseh.com

Deep Learning-based Automated Detection of Facial Surgeries Using HDA Plastic Surgery Database

Amirhossein Rahmani[1*], Saman Ghedasati[2]

[1] Department of Plastic and Reconstructive Surgery, Iranshahr University of Medical Sciences, Iranshahr, Iran (Amirho.rahmani@gmail.com)

[2] Islamic Azad University, Tehran Medical Branch, Tehran, Iran

Received date: July 30, 2024; Accepted date: November 21, 2024

Abstract

Plastic surgery has gained significant traction in modern medicine, with procedures like facial bone correction and nose correction becoming increasingly popular. These surgeries often result in substantial changes to facial features, challenging traditional methods of image analysis and recognition. This study leverages the HDA Plastic Surgery Face Database and state-of-the-art deep learning models—Xception, Vision Transformer (ViT), and Swin Transformer—to classify facial images into five distinct categories: eyebrow correction, eyelid correction, facelifts, facial bones correction, and nose correction. The dataset was preprocessed with image augmentation, normalization, and resizing to enhance model performance. Each model was fine-tuned to capture the subtle variations introduced by different surgeries. Results demonstrate the effectiveness of deep learning in this domain, with Swin Transformer achieving the highest accuracy of 95.5%, precision of 96.9%, and sensitivity of 95.1%.

Keywords*:* Plastic Surgery; Facial Surgery; Vision Transformer; Swin Transformer.

Introduction

Plastic surgery has emerged as a transformative field in modern medicine, enabling individuals to enhance or reconstruct their facial features. It encompasses a wide range of procedures, including cosmetic and reconstructive surgeries, that address both aesthetic and medical needs [1,2]. Among the most common facial surgeries are eyebrow, eyelid, facelift, facial bone, and nose correction. These procedures significantly change facial anatomy, often altering key features that traditional methods of facial analysis and recognition rely on. Analyzing the outcomes of facial surgeries is vital for multiple reasons. Firstly, it provides both patients and surgeons with valuable feedback to evaluate the success of the procedures and whether they meet preoperative goals. Secondly, automated methods can offer a standardized approach to surgical outcome analysis, reducing the subjectivity of manual assessments. This is particularly useful in training environments, where aspiring surgeons need to understand the nuances of various surgical transformations. Imaging analysis in surgery enables precise preoperative planning, intraoperative guidance, and postoperative assessment, enhancing surgical accuracy, patient safety, and outcomes [3,4]. Despite its importance, the analysis of surgical outcomes remains a complex task due to the diversity and subtlety of changes introduced by facial surgeries.

Machine learning (ML), a branch of artificial intelligence, has revolutionized numerous medical applications by offering solutions to intricate problems. ML models have shown significant potential in disease diagnosis [5-7], treatment prediction [8,9], and medical imaging analysis [10-14]. These models excel in detecting patterns and insights from large datasets, often surpassing traditional methods. In medical imaging, for instance, ML has enhanced tasks such as cancer detection [15,16], organ segmentation [17,18], and prediction of treatment outcomes, thereby contributing to personalized medicine and improving patient care [19,20]. Within the realm of machine learning, deep learning has emerged as a game-changer, particularly in the analysis of medical images. Deep learning models can automatically extract hierarchical features from raw data, making them highly effective for image classification, segmentation, and pattern recognition. The adaptability of deep learning methods makes them particularly useful for handling the intricate transformations associated with plastic surgeries, where traditional systems may falter. Deep learning has already found applications in various aspects of plastic surgery, paving the way for advancements in patient care [21], surgical planning [22,23], and postoperative assessments [24]. One notable application is preoperative planning, where models are trained to predict surgical outcomes based on patient anatomy [25,26]. By simulating potential results, deep learning assists surgeons in tailoring procedures to individual needs, improving patient satisfaction, and reducing the likelihood of complications.

Deep learning plays a vital role in evaluating surgical results in postoperative assessments. Models trained to analyze postoperative images can provide quantitative metrics on facial symmetry, proportionality, and aesthetic improvement. Such systems help ensure that surgical outcomes align with the intended goals, offering an unbiased perspective that complements surgeons' manual evaluations. Another critical application of deep learning in plastic surgery is enhancing facial recognition systems [27,28]. Plastic surgery-induced changes can significantly disrupt traditional facial recognition algorithms. Deep learning models, however, are capable of learning invariant features, making them robust to transformations caused by surgeries. This capability is crucial in security and identification systems, ensuring accurate recognition even in cases where facial modifications are substantial. Deep learning also contributes to educational tools for surgical training [29]. By automating the classification and evaluation of surgical images, these models help train surgeons to recognize and understand the outcomes of various procedures. This is particularly useful for developing expertise in complex surgeries, where subtle differences in outcomes are critical for success. Furthermore, deep learning can assist in predicting long-term surgical outcomes [30,31]. By integrating preoperative data with postoperative images, these models can forecast how surgical results will evolve, helping surgeons make informed decisions about the techniques and materials they use.

In this study, we focus on a specific application of deep learning in plastic surgery: the classification of facial images based on the type of surgery performed. Using the HDA Plastic Surgery Face Database, which includes images before and after surgeries, we classify these into five categories:

eyebrow correction, eyelid correction, facelift, facial bones correction, and nose correction. To accomplish this, we employ state-of-the-art deep learning models, including Xception, Swin Transformer, and Vision Transformer (ViT). By automating the classification of facial surgeries, this work demonstrates the potential of deep learning to standardize surgical assessments, enhance medical training, and provide valuable feedback for patients and surgeons alike.

Methods and Materials

Dataset

The study utilized the HDA Plastic Surgery Face Database [32], a specialized dataset curated to analyze the effects of plastic surgery on facial features. This dataset consists of 1,278 facial images captured both before and after surgical procedures, providing a comprehensive resource for studying surgery-induced transformations. The database includes images representing five categories of common facial surgeries:

- *Eyebrow Correction*: Adjustments to eyebrow shape and positioning.
- *Eyelid Correction:* Surgeries addressing eyelid structure, such as blepharoplasty.
- *Facelift*: Procedures targeting overall facial rejuvenation and wrinkle reduction.
- *Facial Bones Correction*: Structural modifications to the bones of the face, often aimed at reshaping or reconstructing features.
- *Nose Correction*: Procedures such as rhinoplasty to enhance or reconstruct the nasal structure.

The dataset ensures a balanced representation of the five surgery categories, enabling robust model training and evaluation. Each image is labeled according to its respective surgery type, facilitating supervised learning tasks. The inclusion of diverse facial poses and lighting conditions enhances the dataset's suitability for training models to generalize well to real-world scenarios. To prepare the dataset for deep learning models, preprocessing steps were applied to normalize and augment the images.

Deep Learning Models

Deep learning models can be broadly categorized into convolutional neural network (CNN)-based models and transformer-based models, each with unique advantages for image classification tasks. CNN-based models, like Xception, excel in capturing local patterns and spatial hierarchies through convolutional layers, making them highly effective for tasks requiring detailed feature extraction, such as identifying subtle changes in facial features post-surgery. Transformer-based models, including Vision Transformer (ViT) and Swin Transformer, leverage self-attention mechanisms to capture both local and global relationships within images. These architectures are particularly suited for analyzing complex transformations, offering enhanced adaptability and scalability. In this study, we employ Xception alongside ViT and Swin Transformer to classify facial images into five surgery categories, leveraging their complementary strengths.

Xception

The Xception model [33] is a deep convolutional neural network based on depthwise separable convolutions, which decompose a standard convolution operation into two parts: a depthwise convolution and a pointwise convolution. This architectural innovation significantly reduces the number of parameters and computational costs while maintaining or improving the model's ability to learn complex features.

Figure 1 Examples of 5 different plastic surgeries 1st row: eyebrow, 2nd row: eyelid, third row: facelift, fourth row: facial bones, and fifth row: nose correction [32].

Technically, the depthwise convolution applies to a single filter to each input channel, extracting spatial features independently, while the pointwise convolution applies to a 1x1 kernel to combine these spatial features across all channels. By separating spatial and depth operations, Xception effectively captures both fine-grained and higher-level representations of the input image. In this study, the Xception model was pre-trained on ImageNet, providing a strong initial feature representation. The network was fine-tuned on the HDA Plastic Surgery Face Database by replacing the original classification layer with a custom fully connected layer, configured to classify the images into five categories. The final output layer used a softmax activation function to provide probabilities for each class. The model's architecture, with its hierarchical feature extraction capabilities, made it particularly effective at identifying localized changes, such as those seen in eyebrow and eyelid corrections. However, its performance was limited in scenarios requiring a broader contextual understanding, such as facelifts or facial bone corrections.

Vision Transformer (ViT)

The Vision Transformer (ViT) [34] is a novel architecture that applies transformer-based self-attention mechanisms to image classification tasks. Unlike traditional CNNs, ViT processes an image as a sequence of patches, treating each patch as a token similar to words in natural language processing tasks. Each patch is first flattened into a one-dimensional vector and then passed through a linear projection layer, which embeds the patch into a high-dimensional space.

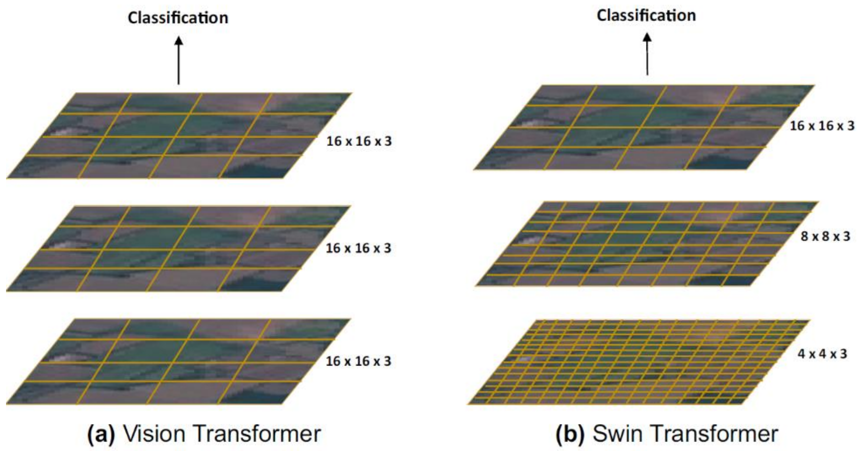

Figure 2 Vision Transformer vs Swin Transformer [35]

The self-attention mechanism in ViT operates by computing relationships between all pairs of patches, capturing both local and global dependencies. This is achieved through the multi-head self-attention (MHSA) mechanism, where multiple attention heads independently learn different feature representations. These features are aggregated to provide a comprehensive understanding of the image. The ViT model used in this study was pre-trained on large-scale datasets and fine-tuned on the HDA Plastic Surgery Face Database. Images were resized and divided into fixed-size patches (e.g., 16x16 pixels), and positional embeddings were added to retain spatial information. The transformer encoder then processed these embeddings, capturing intricate relationships between patches. The final classification layer mapped the learned representations to the five surgery categories. ViT demonstrated remarkable ability in capturing global transformations, making it highly effective for surgeries that involve significant structural changes, such as facial bone correction and facelifts.

Swin Transformer

The Swin Transformer [36] builds on the principles of transformers while addressing their scalability challenges for high-resolution image processing. Unlike ViT, which processes the entire image as a sequence of patches, the Swin Transformer introduces a hierarchical architecture that computes attention within smaller, non-overlapping windows. This window-based attention mechanism enables the model to focus on local regions, significantly reducing computational complexity. A key innovation in the Swin Transformer is the shifted window attention, where the window boundaries shift between layers to ensure connections between adjacent windows. This shifting mechanism allows the model to aggregate local features into a global context as the hierarchy progresses through the network. Additionally, the hierarchical structure progressively reduces image resolution, enabling multi-scale feature learning like CNNs. In this study, the Swin Transformer was pre-trained on large-scale datasets and fine-tuned for the classification task. The model split each image into fixed-size windows, applied attention within each window, and aggregated these features across layers using the shifted window mechanism. The final classification head mapped the learned features to the five surgery categories. Swin Transformer excelled in distinguishing between subtle and overlapping features, such as those present in facelifts and nose corrections. Its ability to balance local feature extraction with global context understanding made it the top-performing model in this study, achieving the highest accuracy, precision, and sensitivity across all categories.

Experiments

Training Process and Configuration

The training process involved fine-tuning the deep learning models—Xception, Vision Transformer (ViT), and Swin Transformer—on the HDA Plastic Surgery Face Database to classify facial images into five distinct categories: eyebrow correction, eyelid correction, facelift, facial bones correction, and nose correction. Each model was initialized with pre-trained weights, leveraging knowledge from large-scale datasets like ImageNet. This approach allowed the models to begin training with a strong feature representation and focus on learning the specific features related to plastic surgery transformations.

The dataset, comprising 1,278 images, was divided into three subsets: 70% for training, 15% for validation, and 15% for testing. Care was taken to ensure a balanced representation of all five categories within each subset, avoiding class imbalance issues that could bias the models during training or evaluation. Preprocessing steps included resizing all images to a standard input size of 224x224 pixels to meet the input requirements of the models. Additionally, pixel values were normalized to the range [0, 1], which helps stabilize and speed up the training process by ensuring that input features have similar scales. To enhance the robustness and generalizability of the models, various data augmentation techniques were applied during training. These included random rotations, horizontal flips, brightness adjustments, and cropping. These augmentations simulated variations in facial poses and lighting conditions, enabling the models to perform well on diverse input scenarios. The training utilized the categorical cross-entropy loss function, suitable for multi-class classification tasks, and the Adam optimizer with an initial learning rate of 0.001. A learning rate scheduler was employed to dynamically adjust the learning rate during training, promoting stable convergence and avoiding overfitting. The models were trained for a maximum of 50 epochs, with early stopping based on validation performance to prevent unnecessary computations and potential overfitting.

The training process was carried out on an NVIDIA GPU with 16GB memory to ensure efficient handling of the computational demands of deep learning. The validation set was used to monitor the models' performance during training, allowing for hyperparameter adjustments and the selection of the best-performing model configuration. After training, the models were evaluated on the test set to determine their final performance in terms of accuracy, precision, sensitivity, and F1 score. This rigorous training process ensured that the models were well-optimized for the task of facial surgery classification.

Results

The classification models—Xception, Vision Transformer (ViT), and Swin Transformer—were evaluated on the HDA Plastic Surgery Face Database, with metrics including accuracy, precision, sensitivity, and F1 score. The results of these evaluations are summarized in Table 1.

Table1. Performance comparison of different deep learning models

Model	Accuracy	Precision	Sensitivity	F1score
Xception	87.3%	88.1%	79.6%	83.63%
ViT	93.2%	94.5%	91.8%	93.13%
Swin	95.5%	96.9%	95.1%	95.99%

The results indicate that the Swin Transformer outperformed both Xception and Vision Transformer in all metrics, achieving an accuracy of 95.5%, a precision of 96.9%, a sensitivity of 95.1%, and an F1 score of 95.99%. Its ability to balance local and global feature extraction using hierarchical and shifted window attention mechanisms likely contributed to its superior performance. The model excelled particularly in distinguishing overlapping features, such as those in facelifts and nose corrections, where both subtle and broad structural changes occur. The Vision Transformer also performed strongly, with an accuracy of 93.2%, precision of 94.5%, sensitivity of 91.8%, and an F1 score of 93.13%. Its self-attention mechanism, which models global dependencies across image patches, proved effective in analyzing surgeries involving large-scale transformations, such as facial bone correction and facelifts. However, its performance was slightly lower than the Swin Transformer, possibly due to its lack of hierarchical feature aggregation. The Xception model, while achieving a respectable accuracy of 87.3%, had the lowest precision (88.1%), sensitivity (79.6%), and F1 score (83.63%) among the three models. As a CNN-based model, Xception excelled in capturing localized changes, making it suitable for surgeries like eyebrow and eyelid corrections. However, its performance was limited in scenarios requiring a broader contextual understanding, such as facelifts and facial bone corrections, where global feature extraction is crucial. The confusion matrices for each model revealed key insights into their classification capabilities: The Swin Transformer showed minimal misclassifications across all five surgery categories, particularly excelling in distinguishing facelifts from other procedures. The Vision Transformer demonstrated strong classification performance but showed slight confusion between closely related surgeries like nose corrections and facelifts. The Xception model struggled more with overlapping features, leading to increased misclassifications in surgeries with broader structural changes.

In summary, the hierarchical architecture and shifted window attention of the Swin Transformer gave it an edge in handling both local and global transformations. This makes it particularly effective for tasks involving complex image relationships, such as facial surgery classification. Furthermore, both transformer-based models (ViT and Swin Transformer) significantly outperformed the CNN-based Xception model, highlighting the importance of self-attention mechanisms in capturing global dependencies. While Xception performed comparatively lower overall, its localized feature extraction strengths make it a viable option for tasks focused on subtle, localized changes.

Limitations and Future Works

While the results of this study highlight the effectiveness of deep learning models, particularly transformer-based architectures, in classifying facial surgeries, certain challenges and areas for improvement remain. Addressing these limitations and expanding the scope of the research can further refine the utility of these models in clinical applications. Below, we outline the limitations:

- Dataset Size: The HDA Plastic Surgery Face Database contains 1,278 images, which, while valuable, may not fully capture the variability in facial features and surgical outcomes across diverse populations. A larger dataset with greater demographic diversity (e.g., age, ethnicity, and gender) would enhance the model's generalizability.

- Class Imbalance: Although efforts were made to balance the dataset, certain surgery categories may have fewer samples, potentially impacting the model's ability to perform equally well across all classes.

- Limited Surgery Types: The study focuses on five common facial surgeries, but the inclusion of additional types, such as jaw surgeries, lip augmentations, or skin resurfacing, could provide a more comprehensive analysis of surgical transformations.

- Pose and Lighting Variations: While data augmentation partially addressed variability in pose and lighting, real-world scenarios might include more extreme conditions that were not fully simulated in the training set.

- Overfitting Risk in Pre-trained Models: Fine-tuning pre-trained models on relatively small datasets risks overfitting, despite regularization techniques such as dropout and data augmentation.

- Limited Focus on Explainability: The study primarily focuses on classification performance without delving into explainability techniques, such as Grad-CAM or SHAP, which could provide insights into the specific features the models rely on for predictions.

This study demonstrates the potential of deep learning in facial surgery classification but highlights several areas for improvement. Expanding the dataset to include more diverse demographics and additional surgery types would enhance model robustness and applicability. While ImageNet pre-trained models performed well, pretraining on surgery-specific datasets could improve feature extraction. Incorporating multimodal data, such as patient metadata, alongside images, could provide more holistic insights. Future work should also explore explainability techniques like Grad-CAM to better understand the features influencing predictions, making models more interpretable for clinical use. Testing in real-world environments would validate the models’ practical utility and robustness. Beyond classification, developing models to predict surgical outcomes and applying deep learning to other areas of plastic surgery, such as reconstructive or body surgeries, could broaden its impact. Addressing class imbalance through synthetic data generation and advanced techniques like GANs could further improve performance in underrepresented categories.

Conclusion

This study demonstrates the effectiveness of deep learning models in classifying facial images into five common types of plastic surgeries—eyebrow correction, eyelid correction, facelift, facial bones correction, and nose correction—using the HDA Plastic Surgery Face Database. Among the models tested, the Swin Transformer achieved the highest accuracy (95.5%), precision (96.9%), and F1 score (95.99%), outperforming the Vision Transformer (ViT) and Xception. The results highlight the capability of transformer-based architectures to handle both local and global transformations effectively, making them particularly suitable for complex image classification tasks in plastic surgery. Future work should focus on addressing these limitations and expanding the scope of the research to other areas of plastic surgery, including reconstructive and body procedures. By integrating multimodal data and deploying models in real-world environments, future studies can further refine the role of deep learning in plastic surgery, ultimately contributing to improved surgical practices and patient outcomes.

Conflict of Interest

The authors imply no conflict of interest.

References

[1] Branford OA, Kamali P, Rohrich RJ, Song DH, Mallucci P, Liu DZ, Lang D, Sun K, Stubican M, Lin SJ. # PlasticSurgery. Plastic and reconstructive surgery. 2016 Dec 1;138(6):1354-65.

[2] King TW. Plastic surgery. Building a Clinical Practice. 2020:203-14.

[3] Rahmani A, Norouzi F, Machado BL, Ghasemi F. Psychiatric Neurosurgery with Advanced Imaging and Deep Brain Stimulation Techniques. International Research in Medical and Health Sciences. 2024 Nov 1;7(5):63-74.

[4] Chuang J, Barnes C, Wong BJ. Overview of facial plastic surgery and current developments. The Surgery Journal. 2016 Mar;2(01):e17-28.

[5] Abbasi H, Afrazeh F, Ghasemi Y, Ghasemi F. A Shallow Review of Artificial Intelligence Applications in Brain Disease: Stroke, Alzheimer's, and Aneurysm. International Journal of Applied Data Science in Engineering and Health. 2024 Oct 5;1(2):32-43.

[6] Kumar Y, Koul A, Singla R, Ijaz MF. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. Journal of ambient intelligence and humanized computing. 2023 Jul;14(7):8459-86.

[7] Ghaffar Nia N, Kaplanoglu E, Nasab A. Evaluation of artificial intelligence techniques in disease diagnosis and prediction. Discover Artificial Intelligence. 2023 Jan 30;3(1):5.

[8] Zhang B, Shi H, Wang H. Machine learning and AI in cancer prognosis, prediction, and treatment selection: a critical approach. Journal of multidisciplinary healthcare. 2023 Dec 31:1779-91.

[9] Liao J, Li X, Gan Y, Han S, Rong P, Wang W, Li W, Zhou L. Artificial intelligence assists precision medicine in cancer treatment. Frontiers in oncology. 2023 Jan 4;12:998222.

[10] Ghasemi F, Minoo S. Dental Imaging Analysis with Artificial Intelligence. Available at SSRN 5008235. 2024 Oct 21.

[11] Kebaili A, Lapuyade-Lahorgue J, Ruan S. Deep learning approaches for data augmentation in medical imaging: a review. Journal of Imaging. 2023 Apr 13;9(4):81.

[12] Rezaei M, Molani S, Firoozeh N, Abbasi H, Vahedifard F, Orouskhani M. Evolving Tsukamoto Neuro Fuzzy Model for Multiclass Covid 19 Classification with Chest X Ray Images. arXiv preprint arXiv:2305.10421. 2023 May 17.

[13] Radak M, Lafta HY, Fallahi H. Machine learning and deep learning techniques for breast cancer diagnosis and classification: a comprehensive review of medical imaging studies. Journal of Cancer Research and Clinical Oncology. 2023 Sep;149(12):10473-91.

[14] Orouskhani M, Zhu C, Rostamian S, Zadeh FS, Shafiei M, Orouskhani Y. Alzheimer's disease detection from structural MRI using conditional deep triplet network. Neuroscience Informatics. 2022 Dec 1;2(4):100066.

[15] Adam R, Dell’Aquila K, Hodges L, Maldjian T, Duong TQ. Deep learning applications to breast cancer detection by magnetic resonance imaging: a literature review. Breast Cancer Research. 2023 Jul 24;25(1):87.

[16] Bosma JS, Saha A, Hosseinzadeh M, Slootweg I, de Rooij M, Huisman H. Semisupervised learning with report-guided pseudo labels for deep learning–based prostate cancer detection using biparametric MRI. Radiology: Artificial Intelligence. 2023 Jul 26;5(5):e230031.

[17] Conze PH, Andrade-Miranda G, Singh VK, Jaouen V, Visvikis D. Current and emerging trends in medical image segmentation with deep learning. IEEE Transactions on Radiation and Plasma Medical Sciences. 2023 Apr 10;7(6):545-69.

[18] Jiang H, Diao Z, Shi T, Zhou Y, Wang F, Hu W, Zhu X, Luo S, Tong G, Yao YD. A review of deep learning-based multiple-lesion recognition from medical images: classification, detection and segmentation. Computers in Biology and Medicine. 2023 May 1;157:106726.

[19] Ahmad A, Tariq A, Hussain HK, Gill AY. Revolutionizing Healthcare: How Deep Learning is poised to Change the Landscape of Medical Diagnosis and Treatment. Journal of Computer Networks, Architecture and High Performance Computing. 2023 Jul 11;5(2):458-71.

[20] Rana MS, Shuford J. AI in Healthcare: Transforming Patient Care through Predictive Analytics and Decision Support Systems. Journal of Artificial Intelligence General Science (JAIGS) ISSN: 3006-4023. 2024 Jan 22;1(1).

[21] Choi E, Leonard KW, Jassal JS, Levin AM, Ramachandra V, Jones LR. Artificial intelligence in facial plastic surgery: a review of current applications, future applications, and ethical considerations. Facial Plastic Surgery. 2023 Jun 23.

[22] Morris MX, Rajesh A, Asaad M, Hassan A, Saadoun R, Butler CE. Deep learning applications in surgery: Current uses and future directions. The American Surgeon. 2023 Jan;89(1):36-42.

[23] Kato RM, Parizotto JD, Oliveira PH, Gonçalves JR. Artificial intelligence in orthognathic surgery–a narrative review of surgical digital tools and 3D orthognathic surgical planning. Journal of the California Dental Association. 2023 Dec 31;51(1):2202444.

[24] Atiyeh B, Emsieh S, Hakim C, Chalhoub R. A narrative review of artificial intelligence (AI) for objective assessment of aesthetic endpoints in plastic surgery. Aesthetic Plastic Surgery. 2023 Dec;47(6):2862-73.

[25] Zhang H, Zhao J, Farzan R, Alizadeh Otaghvar H. Risk predictions of surgical wound complications based on a machine learning algorithm: A systematic review. International Wound Journal. 2024 Jan;21(1):e14665.

[26] Elfanagely O, Toyoda Y, Othman S, Mellia JA, Basta M, Liu T, Kording K, Ungar L, Fischer JP. Machine learning

and surgical outcomes prediction: a systematic review. Journal of Surgical Research. 2021 Aug 1;264:346-61.

[27] Boonipat T, Asaad M, Lin J, Glass GE, Mardini S, Stotland M. Using artificial intelligence to measure facial expression following facial reanimation surgery. Plastic and reconstructive surgery. 2020 Nov 1;146(5):1147-50.

[28] Murphy DC, Saleh DB. Artificial intelligence in plastic surgery: what is it? Where are we now? What is on the horizon?. The Annals of The Royal College of Surgeons of England. 2020 Oct;102(8):577-80.

[29] Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The Virtual Operative Assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. PloS one. 2020 Feb 27;15(2):e0229596.

[30] Tseng PY, Chen YT, Wang CH, Chiu KM, Peng YS, Hsu SP, Chen KL, Yang CY, Lee OK. Prediction of the development of acute kidney injury following cardiac surgery by machine learning. Critical care. 2020 Dec;24:1-3.

[31] Khalid S, Goldenberg M, Grantcharov T, Taati B, Rudzicz F. Evaluation of deep learning models for identifying surgical actions and measuring performance. JAMA network open. 2020 Mar 2;3(3):e201664-.

[32] Rathgeb C, Dogan D, Stockhardt F, De Marsico M, Busch C. Plastic surgery: An obstacle for deep face recognition?. InProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops 2020 (pp. 806-807).

[33] Chollet F. Xception: Deep learning with depthwise separable convolutions. InProceedings of the IEEE conference on computer vision and pattern recognition 2017 (pp. 1251-1258).

[34] Dosovitskiy A. An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929. 2020.

[35] Khan M, Hanan A, Kenzhebay M, Gazzea M, Arghandeh R. Transformer-based land use and land cover classification with explainability using satellite imagery. Scientific Reports. 2024 Jul 20;14(1):16744.

[36] Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S, Guo B. Swin transformer: Hierarchical vision transformer using shifted windows. InProceedings of the IEEE/CVF international conference on computer vision 2021 (pp. 10012-10022).