IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.Doi Number

Tailoring Patient-Specific Cranial Implants for Bone Reconstruction via End-to-End Deep Learning Image-to-Print Approach

Mahmoud Salem[1], Omar Wael[2], Moataz M. Attallah[3], Ahmed Elkaseer[4,5]

[1]Institue for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology (KIT), Eggenstein-Leopoldshafen, 76344, Germany
[2]School of Business, American University in Cairo (AUC), Cairo, New Cairo, 11835, Egypt
[3]Department of Materials, Loughborough University, Epinal Way, Loughborough LE11 3TU, UK
[4]Faculty of Engineering, British University in Egypt (BUE), Cairo, El Sherouk City, Suez Desert Road, 11837, Egypt
[5]Faculty of Engineering, Port Said University, Port Fuad, 42526, Egypt

Corresponding author: Mahmoud Salem (e-mail: mahmoud.salem@kit.edu).

This paper is based upon work supported by the Science and Technology Development Fund Authority (STDF) under EGYPT-UK Cooperation: Research Environment Grant (Project ID: 49831), and in part by the British Council Research Environments Grant: Forging Ahead Together: Strengthening Capacity and Mutual Collaboration in AI & 3D Health Tech between Egypt and the UK (RE2023-096). The authors would like to acknowledge the support provided by the KIT-Publication Fund of the Karlsruhe Institute of Technology.

ABSTRACT Timely, patient-specific cranial implants are critical for restoring skull integrity after trauma, accidents or surgery, yet off-the-shelf devices arrive in coarse, fixed size increments that often fail to match irregular defects. Surgeons are therefore forced to open and try multiple implants intra-operatively, adding infection risk, operative time, and likelihood of early implant failure. This paper presents an end-to-end workflow that predicts "print-ready" cranial implants directly from defective CT volumes and validates them through physical prototyping. The pipeline couples a five-stage volumetric preprocessing routine with a systematic exploration of neural architectures. A baseline and two deeper 3-D U-Nets are benchmarked alongside the proposed version 3 model, a channel-rebalanced U-Net that shifts capacity from shallow texture filters to boundary-aware decoding paths. Two public datasets, SciData and SkullFix, were merged, and split for training the model. On 40 unseen skulls, the proposed version 3 achieved a mean Dice of 0.901, Boundary Dice of 0.908, and HD95 of 1.52 mm, which is surpassing all alternatives and reducing the number of intra-operative trials predicted by simulation from 3-4 to zero. Bench-top tests confirmed that implants printed directly from the network's STL output seated flush on three morphologically diverse phantoms without CAD edits or sanding, demonstrating true "push-button" manufacturability. These results show that a carefully optimized architecture, paired with robust pre-processing, can supply size-perfect implants using modest data and commodity hardware.

INDEX TERMS Computer Vision, Convolutional Neural Networks (CNNs), Custom Cranial Implants, U-Net Structure, Image Processing, 3D Reconstruction, 3D Printing.

I. INTRODUCTION

In the global effort to build safe, healthy societies, every country now treats healthcare as a central pillar which is recognizing that human capital reinforces national progress. A key objective is universal access to high-quality services, such as dependable medical replacements and implants. Yet reliance on imported implants introduces several hurdles. Long procurement lead times and shipping delays can postpone operations for weeks or months [1], while limited size ranges often force surgeons to open and trial multiple devices during surgery [2, 3]. That trial-and-error approach lengthens anesthesia time [4], raises infection risk with each additional package opened, and increases blood loss [5] and tissue trauma from repeated fitting, it can also culminate in a "close-enough" implant that fails early and requires revision [6].

The afore-mentioned challenges are increasing because of surging global needs linked to diabetes, which can lead to extremity amputations, road-traffic injuries, and war injuries. These factors highlight the crucial need for on-site manufacturing capacity that can provide patient-custom implants tailored specifically to each individual, without mis-matched sizes or significant delivery delays.

Recent high-tech advances, such as artificial intelligence (AI), computer vision, and additive manufacturing (AM), can be integrated and utilized to develop these custom implants. The power of AI, especially in computer-vision algorithms, allows the precise anatomical data captured from computed tomography (CT) scans to be interpreted so the structure and features of the implant can be estimated. This approach replaces traditional computer-aided design (CAD) and computational-modeling tools, which require long operator times to create a design. In addition, AM, also called 3-D printing, enables the production of complex patient-specific implants in a short time and at low cost. Both of these innovative methods have a significant positive effect on healthcare in orthopedics by improving treatment outcomes.

In this paper, the authors focus on cranial defects as a crucial part of orthopedics. Cranial defects usually arise from trauma, such as road accidents; surgery, such as major brain operations; or congenital conditions. These defects can affect brain protection, intracranial-pressure regulation, and cerebrospinal-fluid dynamics [7]. Therefore, custom cranial implants are essential for restoring both the functional and aesthetic aspects of the human skull[8-10].

Traditional methods of implant design and production, often involving third-party suppliers, are time-consuming and expensive. The need for patient-specific implants that precisely fit the defect area while restoring natural cranial aesthetics has driven research towards more efficient and cost-effective solutions [11].

Recent advancements in deep learning [12] have shown great potential for automating the design of custom cranial implants. These approaches aim to generate complete cranial models from defective ones, significantly reducing the time and expertise required for implant design [13]. Some approaches utilize deep learning architectures to recognize the shape of the implant and refine it, demonstrating promising results in both synthetic and clinical cases. There is significant progress in deep learning architectural designs for cranial defect reconstruction. The modification of U-Net [14, 15] architectures to be developed for volumetric data of skull, that carry out good results in public data set, homemade/in public dataset and augmented dataset [16, 17]. Also, the utilizing of deep learning to combine defect reconstruction with iterative procedures to improve implant geometry. In Addition, Symmetry-enforcing networks that utilize the natural symmetry of the skull to enhance reconstruction accuracy was utilized to the estimate the implant [18].

Despite these distinguished progress in this field still there are several challenges persist. Limited datasets remain a primary concern which owing to privacy regulations and the rarity of certain defect types, which restrict the availability of patient data [19, 20]. Additionally, rendering synthetic cranial defects and processing high-resolution 3D data necessitates substantial computational resources. Furthermore, there is a growing demand for lightweight models that preserve high accuracy while delivering results within a short runtime. To address these issues, on-going research directions involve exploring efficient architectures that can operate on resource-constrained devices without compromising accuracy.

Cranial reconstruction still faces three intertwined hurdles. First, the scarcity and heterogeneity of publicly available CT data deprive deep-learning models of the representative examples they need. Second, the labor-intensive CAD clean-up that typically sits between segmentation and fabrication slows the workflow. Third, the persistent difficulty of producing truly size-perfect implants from scant clinical data demands a network that can maintain high accuracy and generalize well despite limited training samples.

This paper tackles all three challenges head-on through two main objectives. Objective one is to enhance the U-Net architecture by introducing a channel-rebalanced design, deliberately thinned in its earliest layers and rewired in its decoder, which delivers millimeter-scale accuracy and robust generalization even on a few hundred volumetric scans. Objective two is to evaluate different modified architectures to identify the most effective modification will led to enhancement of the neural neatwork model and to integrate the best-performing one into a fully automated "image-to-print" pipeline. Anchoring this architecture is a fully automated "image-to-print" pipeline that begins with rigorous five-stage preprocessing (bone-window thresholding, tight cropping, isotropic resampling, dimension standardization, and intensity normalization), proceeds to back-of-skull extraction and direct implant prediction, applies lightweight topology optimization, and culminates in the direct 3-D printing of a size-perfect cranial implant, completely eliminating intermediate CAD edits and manual tuning. Polymer prototypes fabricated via fused-filament printing already seat flush in skull phantoms within a ±1 mm clinical tolerance, demonstrating a genuine push-button solution that paves the way for on-demand metal printing in future surgical workflows.

This paper is organized as follows; after this introduction, the background section presents the theoretical foundations of custom implant estimation and the possible methods for achieving it. Next, the overall workflow section describes the structure of the proposed end-to-end solution. This is followed by the proposed computer vision approach section which includes the proposed design of our custom U-Net architecture. The results of this proposed structure are then presented and discussed. Finally, the paper concludes with a summary and suggestions for future work.

II. BACKGROUND

Computer vision is a field of AI that focuses on enabling computers or machines to recognize and understand visual scenes in a human-like manner [21]. Among its numerous tasks, the most prominent are object detection, object

segmentation, and instance segmentation, all of which have a positive impact on medical engineering [22, 23].

Traditional computer-vision methods rely on hand-crafted features developed by a feature engineer; these features are then fed into rule-based algorithms for image analysis [24]. The performance of such methods often degrades when they must deal with complex or irregular structures, such as images or scans of bone defects [25].

In recent years, deep learning (DL) has enabled modern computer-vision approaches, most particularly Convolutional Neural Networks (CNNs), that interpret visual data without the need for hand-crafted features or rule-based logic [26]. CNNs consist of multiple layers of learnable filters, called kernels, that automatically detect and recognize image features. Each layer extracts progressively higher-level abstractions, moving from simple patterns such as edges and corners to more complex shapes [27]. This hierarchical feature understanding makes CNN models highly effective in medical image analysis, where fine-grained accuracy is crucial and anatomical variability is common [26].

CNN models can learn to distinguish structural patterns that indicate intact bone from those that reveal broken or missing bone, thereby addressing the challenge of recognizing absent features, such as bone fragments, in skull defects. By inputting volumetric data, such as CT scans, into specialized network architectures, such as 3D U-Nets [17], these models learn to detect anomalies and predict the shapes that would naturally fill the gaps. This process generally involves two major steps. Firstly, is feature extraction and localization, where the neural network identifies critical regions in the volumetric data that deviate from expected anatomical structures then effectively localize the missing or deformed segments. Secondly is predictive reconstruction, where once the missing features are identified the network generates a segmentation mask or shape prediction approximating the anatomical structure in the area of the defect. In cranial reconstruction tasks, this predicted mask can then be further refined and converted into a 3D model for implant design and manufacturing [13, 16, 28, 29].

Such data-driven approaches hold several advantages over traditional methods. Instead of relying on manually designed features, they reduce the reliance on manual feature engineering and allow the network to learn directly from examples of healthy anatomy and corresponding defect patterns [30, 31]. These automated approaches are important in medical applications where large volumes of patient imaging must be analyzed accurately and quickly [30]. However, best practice for developing a CNN-based algorithm that identifies cracked areas and estimates the features of a skull custom implant requires a comprehensive workflow [13]. This workflow starts with data acquisition and preprocessing to ensure that image sizes and resolutions are suitable for model training [31]. Model training typically demands significant computational resources, usually GPU-based. Testing is then carried out against ground-truth data to assess model accuracy and generalization [31]. It is also essential to verify that the dataset represents real-world variability. Furthermore, incorporating advanced data-augmentation techniques or synthetic data generation for rare or extreme defects can improve model generalization, enabling more accurate recognition of missing features across diverse patient cases [32, 33].

In the literature, several cranial-reconstruction methods exist for repairing skull defects, each with distinct pros and cons, to the best of the authors' knowledge. For example, manual reconstruction relies on a clinician's expertise to sculpt the missing skull region using computer-aided design tools. This approach is time-consuming and heavily dependent on practitioner skill, but it can yield highly accurate and customized designs [34]. Interpolation and extrapolation techniques, which analyze skull shape and curvature to estimate the missing region, offer a more automated yet less adaptable option, making them unsuitable for complex cases or congenital conditions. To the best of the authors' knowledge, these methods work best for small or relatively simple defects, but their accuracy decreases in more complex or irregular cases [29, 35].

In addition, anatomical-template or symmetry-based reconstruction methods can be used to estimate the features of a custom implant. These approaches rely on standard skull templates or on mirroring the intact side of the skull to reconstruct the missing region [18]. They perform well when symmetrical references are available, such as in unilateral defects, but they are limited in bilateral or more complex, multilaterally affected scenarios. Furthermore, statistical shape modeling (SSM) is another powerful method, particularly useful for reconstructing irregular or large cranial defects, but its effectiveness hinges on the availability of extensive, diverse anatomical datasets [36, 37]. Recent strides in AI, specifically convolutional neural networks (CNNs) and generative adversarial networks (GANs), have shown remarkable capabilities in predicting missing cranial regions. By analyzing extensive collections of skull images, these models learn nuanced geometric patterns, achieving high accuracy even in complex cases. However, the success of AI-driven methods is contingent on enough, high-quality, and domain-specific datasets [16-18].

Transfer learning offers a means of adapting pre-trained models to specialized tasks, often reducing training time and data requirements. Models such as MobileNet [38] or U-net [14, 18] typically trained on extensive datasets such as ImageNet [39], can be repurposed for tasks such object recognition. However, in cranial reconstruction, where anatomical accuracy and domain-specific knowledge are paramount may fail to capture the intricacies of skull geometry [36]. Consequently, this paper refrained from implementing transfer learning, choosing instead to train a custom network from the ground up. This native training allowed for targeted architectural refinements, data augmentation strategies, and loss functions specifically aligned with cranial defect

reconstruction requirements. Although transfer learning can expedite prototyping in contexts where large public datasets exist, the specificity and precision demanded by this paper justified the additional time and resource investment in developing a specialized model.

III. METHODOLOGY

In this paper, the authors present a comprehensive workflow, illustrated in Figure 1. First, the defective skull is imaged with CT, producing cross-sectional data that capture both the defect and the surrounding anatomy. These images are then pre-processed so that subsequent stages can interpret the defect accurately. The AI model estimates the implant geometry, after which a CAD file is generated and exported in standard AM formats such as OBJ or STL. Before fabrication, the file undergoes topology optimization to yield an implant that closely matches the patient's anatomy while maintaining structural robustness. Finally, the optimized file is sent to an AM system capable of fabricating high-precision, biocompatible implants. In this study, the authors first produced polymer prototypes using fused-filament fabrication (FFF) for rapid, low-cost testing; final metal production will follow in future work.

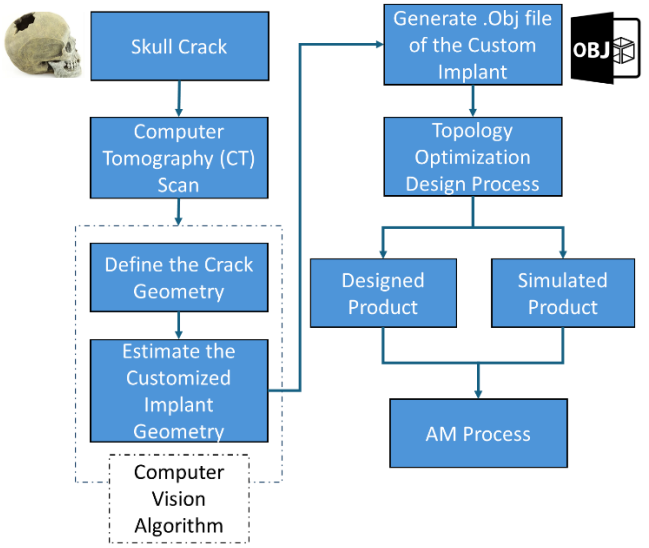

FIGURE 1. Workflow of the proposed method for estimating, designing, and manufacturing a patient-custom skull implant.

IV. PROPOSED COMPUTER VISION APPROACH

A. PROPOSED NEURAL NETWORK STRUCTURE

The U-Net architecture is a type of CNN designed for biomedical image-segmentation tasks [14]. Its characteristic "U-shaped" structure consists of a contracting path, or encoder, and an expansive path, or decoder, as illustrated in Figure 2. The encoder and decoder work together to capture the spatial context of an image and produce an accurate segmented output.

The contracting path is responsible for extracting high-level features through multiple convolutional layers, each followed by a rectified linear unit (ReLU) activation [40] and max-pooling operations. This process progressively reduces the spatial dimensions of the input image while retaining critical feature information.

In the expansive path, the network restores the spatial resolution of the feature maps through transposed convolutions (up-convolutions), which increase the spatial dimensions of the data. A defining feature of U-Net is its skip connections, which directly link corresponding layers in the encoder and decoder. These connections ensure that fine-grained spatial information from the encoder is preserved and integrated with high-level contextual information in the decoder. This methodology enhances model's ability to produce accurate segmentation results.

The encoder and decoder are bridged together by a bottleneck layer located at the center of the U-Net architecture. This layer consists of convolutional operations with a large number of filters, allowing it to capture comprehensive and abstract features of the input image.

At the end of the U-Net, a final output layer with a 1×1 convolution maps the learned features to the desired output classes. For binary segmentation tasks, the model applies a sigmoid activation function, whereas for multi-class segmentation it uses a SoftMax activation [41, 42].

To the best of the authors' knowledge, U-Net offers several advantages: it is an end-to-end architecture that processes raw input images and efficiently generates pixel-wise segmentation maps; and its skip-connection structure enables it to produce high-resolution segmentation results with a high degree of precision while reducing overfitting. These features make U-Net adaptable to a wide range of tasks beyond its original biomedical context [14].

Due to the success of the U-Net network, several extensions have been developed. For example, the Attention U-Net integrates attention mechanisms to focus on relevant regions in the input image [43]. ResUNet++ [44] is another variant that incorporates residual connections to improve gradient flow and convergence. Additionally, 3D U-Net extends the architecture to process volumetric data, such as CT scans, making it particularly valuable for three-dimensional image-analysis tasks [15].

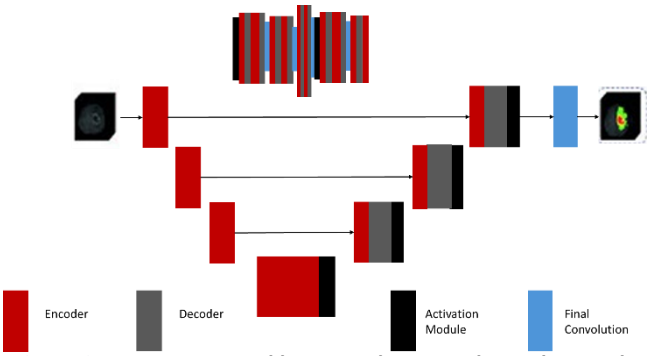

FIGURE 2. U-net structure with symmetric contracting and expansive paths, connected through skip connections. The encoder captures contextual information by progressively down sampling the input, while the decoder reconstructs the spatial resolution through up sampling, enabling precise segmentation. Skip connections ensure the preservation of fine-grained details by directly transferring features from the encoder to the decoder at corresponding levels. (Modified from [45]).

In this paper, a modified version of U-Net is employed to segment cranial defects and predict implants based on defective skull data. Its ability to balance spatial resolution and contextual understanding made it an ideal choice for this application. The model's robust architecture and flexibility ensured accurate segmentation outputs, critical for designing patient-specific implants.

The proposed modified architecture of U-net, referred to version 3 (V3) in this study, begins with a deliberate thinning of the shallow encoder. Where, the authors provide three versions of architecture. Whereas the reference model starts with a 16-channel stride 3-D convolution, Version 3 structure reduces this width to just 8 channels and removes every residual block in the first two stages. This almost halves both the parameter counts and the memory footprint of the network; its residual block is resized accordingly. These bookkeeping changes eliminate the zero-padding previously required to align mismatched tensors, yielding slightly sharper boundary predictions and a cleaner gradient flow back into the shallow encoder (see Figure 3).

Also, the authors developed two trial architectures, version 1 (V1) and version (V2). In contrast to version 3, these experiments deepened the shallow encoder in search of stronger edge detectors. Specifically, version 1 inserted two full residual blocks ahead of every down-sampling operation in stages 1 and 2, while version 2 retained a single extra block per stage after ablation testing showed diminishing returns for the second block. Although both variants slightly improved the training metrics, they also inflated the model by millions of parameters and exhibited clear signs of over-fitting: the

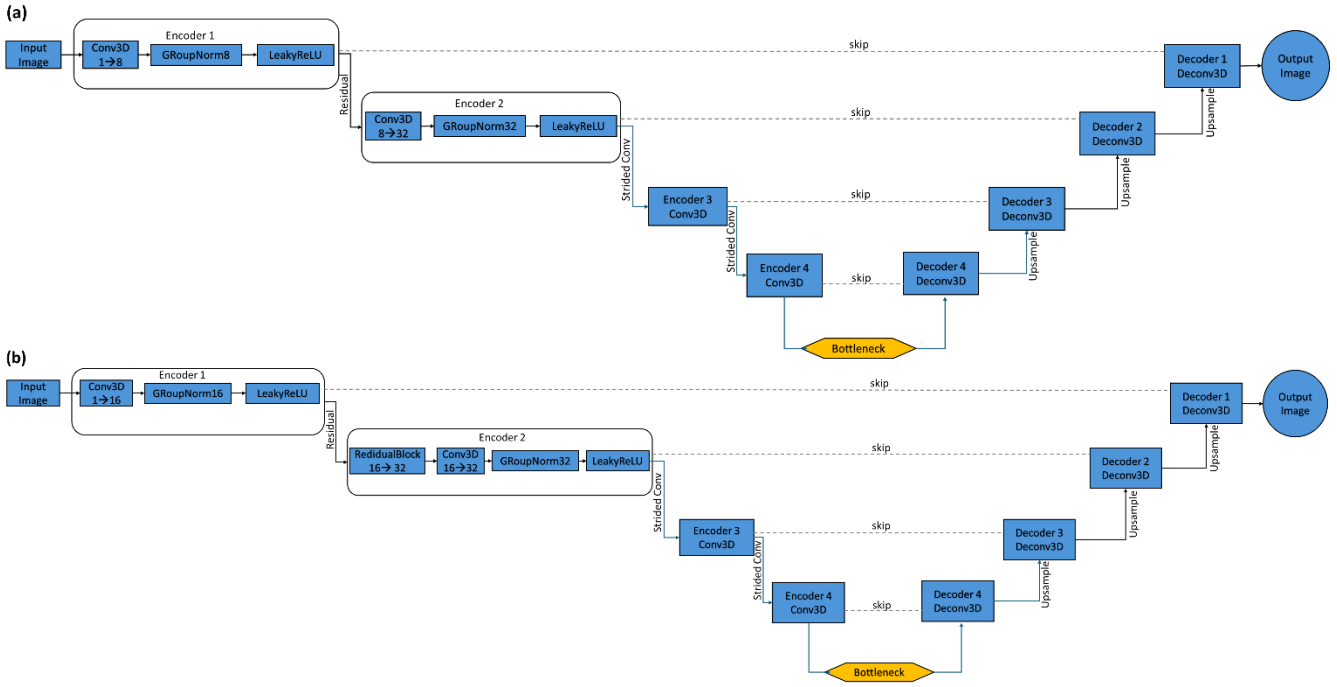

FIGURE 3. (a). The structure of the proposed modification U-Net structure (Version 3). (b) The structure of native U-Net.

low-level path, lowering the risk of early-layer over-fitting on limited skull-implant data while also accelerating training. From the third stage onward, the authors preserve the native design, two residual blocks at 64 channels followed by two at 128 channels, so the network still has ample depth to capture the complex curvature of implants.

In other words, the proposed version 3 architecture reallocates capacity where it trims redundancy in texture-oriented layers and retains the shape-centric context layers, striking a better balance between efficiency and representational power.

A second key modification is the precise rewiring of the decoder to match the new channel plan. Because the first skip feature now arrives with 8 channels instead of 16, the final decoder block processes a 40-channel concatenated tensor (32 + 8) rather than the 48-channel tensor used in the native validation Dice plateaued early and fluctuated with each epoch. The proposed version 3 architecture with its lighter front-end and channel-consistent skip connections, therefore offers a more elegant route to higher accuracy, delivering our best performance metrics to date along with faster convergence and lower computational cost.

B. DATASETS

One significant challenge in this research was the scarcity of publicly available cranial reconstruction datasets. Privacy concerns and restrictions around sharing medical imaging data often limit access to comprehensive collections of high-resolution CT scans of the human skull. Researchers must rely on smaller, fragmented, or synthetic datasets in order to complicate model training and validation. To address this, two datasets were identified and employed. The first one is SciData

which consists of CT scans of 24 healthy skulls where cranial defects have been artificially induced, resulting in 240 paired samples of defective skulls and corresponding implants. This dataset is instrumental for researchers seeking to improve cranial implant design with minimal dependence on manual methods or external manufacturers [46, 47]. The second dataset, SkullFix, was originally introduced during the MICCAI 2020 AutoImplant Challenge. It comprises 100 triplets for training, each triplet consisting of a complete skull, a defective skull, and the corresponding implant, and 110 triplets for evaluation. this dataset provides a richer set of triplets and more clinically realistic geometry. The training partition contains 100 triplets, each consisting of (i) a complete skull, (ii) a corresponding defective skull created by digitally excising bone with realistic surgical margins, and (iii) the "gold-standard" implant model that perfectly restores the anatomy. All volumes are delivered as NIfTI files with 1 mm isotropic spacing and a fixed matrix size of 256 × 256 × 256 voxels, and each sample is accompanied by an STL mesh of the implant for direct 3-D printing. Defects range from small burr-hole–like openings (≈3 cm) to large temporo-parietal craniectomies (>12 cm), covering frontal, parietal, temporal and occipital regions in roughly equal proportions. Age, sex and cranial size are also balanced across the cohort, making SkullFix a de-facto benchmark for volumetric shape learning and implant generation in the craniomaxillofacial community [48, 49]. It provides comprehensive paired data, SkullFix is widely regarded as a benchmark for tasks related to volumetric shape learning.

An important limitation encountered was the inadequacy of basic data augmentation methods. Standard techniques such as rotations, flipping, and scaling are effective for tasks such as general image classification, but they do not create new structural variations crucial for simulating a wide array of cranial defects. Previous studies have shown that such augmentation strategies such as random image rotations or nonlinear deformations can improve performance but are often limited in their ability to emulate real-world variability [50-53]. Despite the current project's time and resource constraints limiting the volume of synthetic data produced, this rendering-based approach promises to significantly enhance the realism and diversity of future datasets. Plans are in place to automate this process to expand the scale of synthetic data generation and improving model robustness in the long term.

C. NEURAL NETWORK TRAINING AND EVALUATION PIPELINE

The implemented pipeline for training and evaluating the U-Net model for automatic cranial implant design is depicted in Figure 3. This workflow integrates multiple datasets and utilizes advanced preprocessing and evaluation techniques to optimize the performance of the model.

The process begins by combining data from two key sources. The SciData dataset which includes 240 samples consisting of cranial defects and their corresponding implants, and the AutoImplant (SkullFix) dataset which provides 100 samples. The combined dataset is split into three groups: 250 samples for the training phase, 50 samples for the validation phase, and 40 samples for the testing phase. The split is designed to balance the data and capture all features systematically, supporting robust model training and evaluation.

During preprocessing, the images are standardized to mitigate inconsistencies before being fed into the U-Net model. This stage includes loading the volumetric data, cropping to focus on the region of interest such as the crack area, resampling to ensure uniform resolution, and padding to maintain consistent input dimensions.

Successful, generalized training of a 3-D U-Net for cranial-defect segmentation depends heavily on the quality and uniformity of the volumetric data. Raw clinical CT scans often exhibit considerable heterogeneity: slice thickness can vary two- to three-fold across scanners, in-plane resolution is rarely isotropic, and the cranial vault may occupy only a fraction of the standard 512 × 512 acquisition field. Left uncorrected, these inconsistencies cause convolutional kernels to learn scanner-specific features, inflate class imbalance, and waste GPU memory on non-informative background voxels, ultimately degrading model accuracy and limiting batch size.

To mitigate these issues, the authors utilized an efficient five stages of pre-processing [16]. First, it gets the CT image where each CT study was first thresholder within a bone window and binarized, yielding a full-field cranial mask that preserved the scanner's native voxel spacing. Although this representation retained all anatomical detail, it also contained large contiguous regions of background that contributed little information but consumed substantial GPU memory. Secondly is to make spatial cropping for the image. Where a tight three-dimensional bounding box was computed around the binarized skull. Cropping reduced the average voxel count by ~75 %, lowering the foreground-to-background imbalance and permitting larger batch sizes without exceeding memory constraints. Third, is to make isotropic resampling. Because slice thickness frequently differed from in-plane resolution such as 0.6 × 0.6 × 1.5 mm, each cropped volume was resampled to an isotropic grid of 1 mm^3 using tri-linear interpolation for images and nearest-neighbor interpolation for labels. This step ensured that a 3 × 3 × 3 convolutional kernels represented the same physical extent in every patient, thereby promoting spatially coherent feature learning across heterogeneous scanners. Fourthly is to make dimension standardization. Which inter-patient variation in head size was normalized by zero-padding each resampled scan to a fixed matrix of 240 × 200 × 240 voxels. Uniform tensor dimensions eliminated run-time shape negotiation, enabling streamlined batch processing and reproducible deployment. Fifthly is intensity normalization and label encoding. Where, voxel intensities were z-scored within the resampled bone window

to mitigate scanner-specific brightness drift. of predictions on the native CT space during inference.

The preprocessed data is then used to train the U-net model, a widely recognized architecture for image segmentation tasks. The model is trained to predict cranial implants from defective skull inputs, utilizing the comprehensive datasets to learn intricate features of cranial geometry.

After training, the model is evaluated with three metrics. First, the Dice Score (DSC) measures the overlap between the predicted and ground-truth regions. Second, the Boundary Dice (BD) assesses boundary accuracy. Finally, the 95th-percentile Hausdorff Distance (HD95) quantifies the maximum distance between points on the predicted and ground-truth boundaries. Together, these metrics give a comprehensive view of segmentation performance: higher DSC and BD values and lower HD95 values indicate better accuracy and alignment.

The predicted outputs are then converted from .nrrd to .obj format to ensure compatibility with the AM machine for physical fabrication of the custom implants.

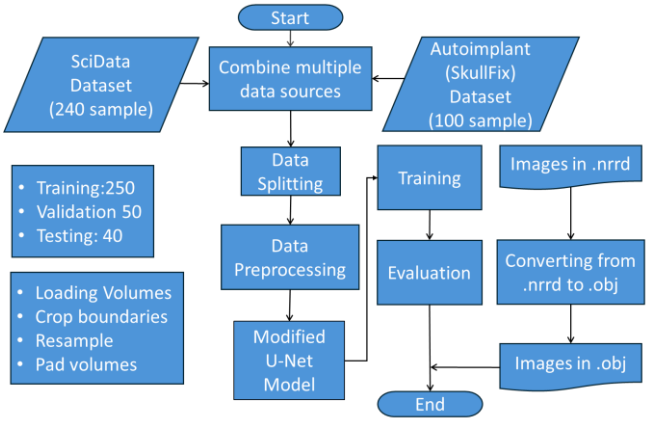

FIGURE 4. Workflow for training and evaluating a U-Net model for automatic cranial implant design.

V. RESULTS AND DISCUSSION

Figure 5 shows the preprocessing results of the dataset which is fed to the neural network. It illustrates the sequential pre-processing steps applied to the binary skull mask and quantifies their cumulative impact on data integrity and computational efficiency. Figure 5 (a) shows the starting from the raw segmentation, an aggressive 3-D bounding-box eliminated the substantial peripheral background, reducing the median volume dimensions from 512 × 512 × 380 voxels to 186 × 154 × 173 voxels while preserving 100 % of foreground voxels (see Figure 5 (b)). This alone cut the per-case memory footprint by roughly an order of magnitude and improved the foreground-to-background ratio that downstream loss functions see during training. Figure 5 (c) shows the resampling to an isotropic 1 mm grid, thereby harmonizing spatial resolution across the cohort. Nearest-neighbor interpolation was applied to the binary masks, while linear interpolation was used for intensity images, confirming that nearest-neighbor is adequate for binary labels at this scale. Padding the resampled crop to a fixed 240 × 200 × 240 array. Figure 5(d) standardized input shapes for the CNN, adding empty margins rather than altering anatomy. Figure 5 (e) shows the final export stage merely writes the tensor to disk with a corrected affine header.

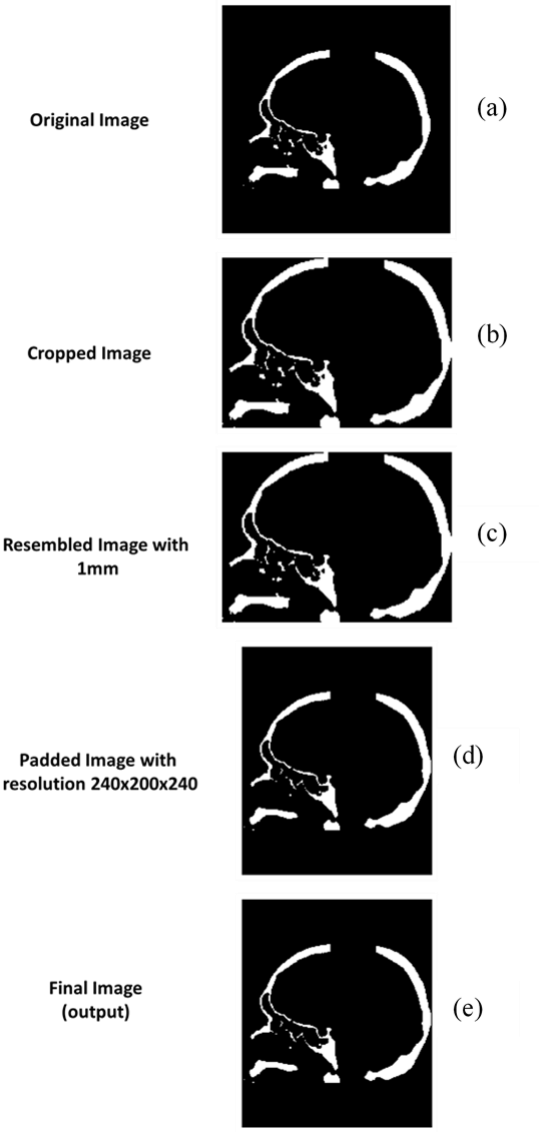

FIGURE 5. Results of preprocessing pipeline.

Figures 6-10 present the training and validation curves for these various models as well as for established architectures from the literature. Throughout these figures, training loss is shown in blue and validation loss in orange, and each model is trained for a total of 300 epochs.

For the case of the native U-Net architecture in Figure 6, the training and validation losses both start relatively high but drop considerably within the first 10 to 15 epochs. The training loss initially hovers around 1.0 while the validation loss spikes near 1.6 to 1.7. As the model begins to learn, the both curves rapidly decline. This decline indicates that the network is quickly adjusting its weights to fit the data. Beyond

approximately epoch 20, the training loss settles around 0.1 to 0.2 and the validation loss tends to stay in the 0.2 to 0.3 range. Although some fluctuations can be observed in the validation curve where there is no severe gap developing, suggesting that while the model may still leave room for improvement, it avoids dramatic overfitting.

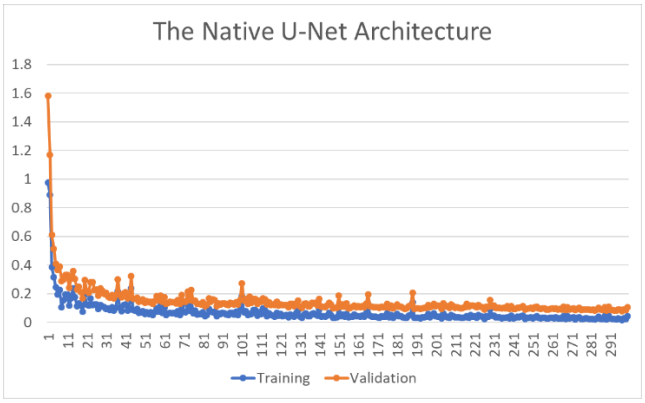

FIGURE 6. Training and validation loss curves over 300 epochs for the native U-Net architecture.

In Figure 7 illustrates the modified U-Net version 1 performance. Its overall trend remains similar in the early epochs with a sharp drop from around 1.6 for the validation loss and around 1.0 for the training loss. Notably, the validation curve settles slightly faster to under 0.2 for a longer span of the training process compared to the native U-Net, suggesting potentially better generalization. While minor spikes appear in the mid-training phases, the model's training curve consistently hovers below the validation curve, reflecting steady convergence without drastic overfitting.

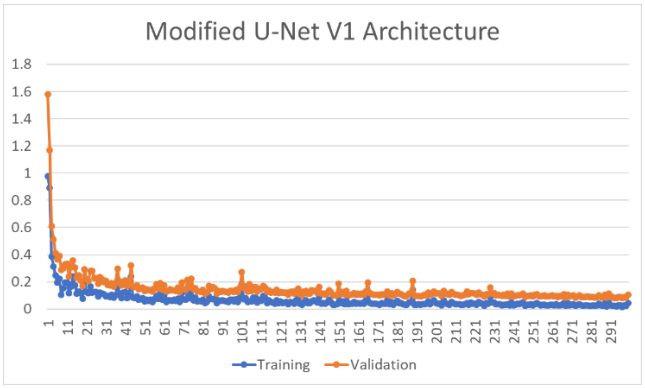

FIGURE 7. Training and validation loss curves over 300 epochs for the Modified U-Net V1 architecture. The plot illustrates how changes in network design influence convergence and generalization performance.

In Figure 8 illustrates the performance of the modified U-Net version 2 model. The model has the same rapid initial descent which is apparent but the validation loss stabilizes more quickly than in the modified U-Net version 1 model. It generally maintaining values around 0.15 to 0.25 for most of the training. The training curve also remains tightly clustered between 0.1 and 0.2 which indicates more consistent performance. This suggests incremental architectural refinements in version 2 model help the network learn robust features early on and retain that advantage throughout training, with relatively less oscillation in the validation loss.

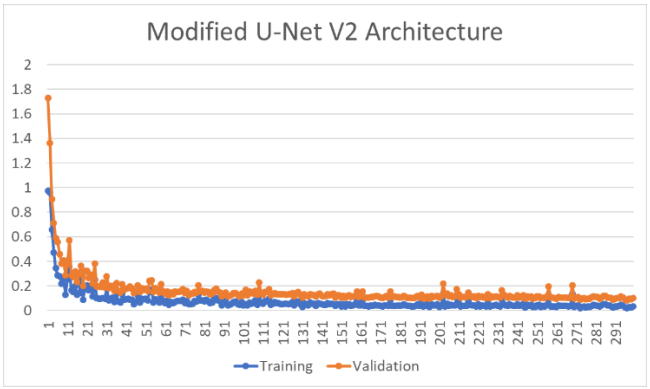

FIGURE 8. Training and validation loss curves over 300 epochs for the Modified U-Net V2 architecture. The plot illustrates how changes in network design influence convergence and generalization performance.

In Figure 9, the proposed modified U-net version 3 model performance is illustrated. Where there is a clear improvement in how swiftly the network converges with the validation loss dropping sharply within the first few epochs and stabilizing to around 0.1 to 0.2 soon thereafter. The training loss also stays consistently lower, generally under 0.15 after the initial descent. This smooth and stable convergence pattern points to effective model regularization and capacity. The noticeable advantage over the previous variants especially in the latter half of training emphasizes how the proposed architectural modifications can yield better generalization and overall performance for the model.

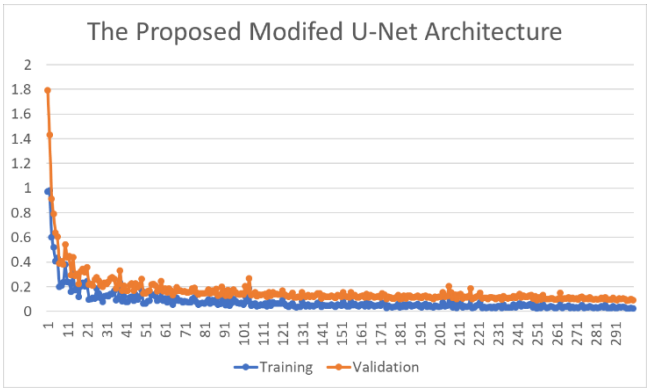

FIGURE 9. Training and validation loss curves over 300 epochs for the Proposed Modified U-Net version 3 model architecture. The plot illustrates how changes in network design influence convergence and generalization performance.

In Figure 10, the pre-trained Cranext model shows both the training and validation curves remain relatively flat throughout the 300 epochs. The training loss stabilizes at roughly 0.6 to 0.7, while the validation loss hovers around 1.3 to 1.4. This persistent gap between training and validation indicates that the model is not reducing its generalization error in response to ongoing training, suggesting that the pretrained weights may not transfer optimally to the current dataset or task. While the model converges quickly in terms of training

loss, the high and stable validation loss reveals limited improvement in out-of-sample performance and highlights the need for either more extensive domain adaptation, additional fine-tuning strategies, or alternative architectures to better fit the data.

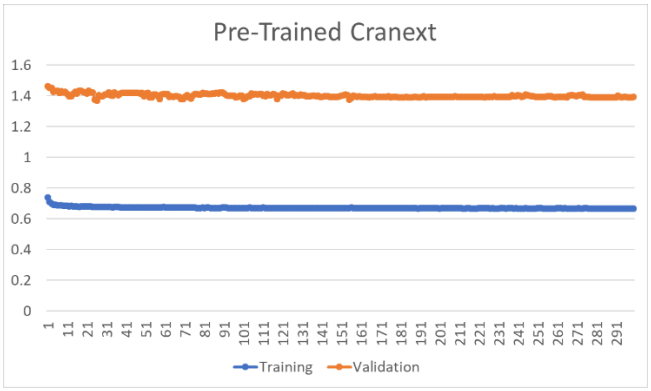

FIGURE 10. Training and validation loss curves over 300 epochs for the Pre-trained Cranext model.

In Figure 11, the native Cranext model, trained from scratch on the present dataset, shows a learning trajectory that contrasts sharply with the pre-trained variant in Figure 10. Both training and validation losses start near 1, fall steeply during the first 15 to 20 epochs, and then continue a gradual, nearly parallel decline for the remainder of the 300 epoch run. By the final epoch, each curve has settled around 0.78 and 0.80, with the separation between them never exceedingly around 0.03 after the initial warm-up phase. This narrow and consistently diminishing gap indicates that the model is improving its generalization performance in tandem with its training accuracy, rather than overfitting or underfitting. Moreover, the steady downward trend, without an early plateau, suggests that the optimizer is still extracting useful features from the data, implying further gains could be achieved with additional epochs or a finer learning-rate schedule. Collectively, these patterns confirm that, for this application, initializing the network with random weights and allowing it to specialize on task-specific images yields markedly better convergence and validation behavior than relying on pretrained weights sourced from a different domain.

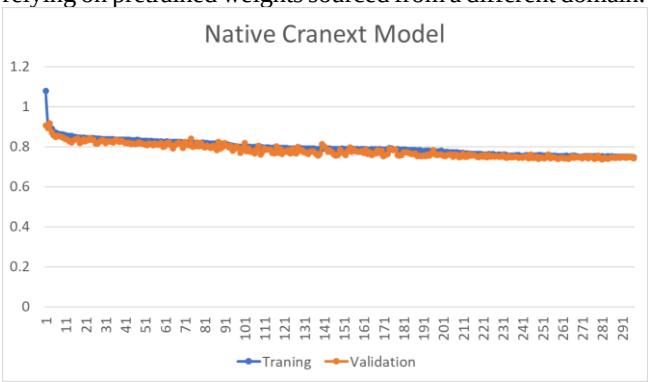

FIGURE 11. Training and validation loss curves over 300 epochs for the native Cranext model.

Table I compares proposed models and literature model across three key performance metrics which are Mean Dice, Mean Boundary Dice, and Mean HD95. The higher Dice score whether overall and boundary dice indicates better overlap with the ground truth segmentation and a lower HD95 reflects fewer extreme boundary errors. From the native U-Net model, the authors observe strong Mean Dice which is 0.8973 and Boundary Dice which is 0.9066 scores, alongside a relatively low HD95 of around 1.54. This baseline demonstrates that standard U-Net already achieves a high level of segmentation accuracy. By contrast, both the Cranext pretrained model and the Cranext native model exhibit lower Dice scores which are 0.8779 and 0.8746, respectively and significantly higher HD95 values which is over 8.3, showing that while their average overlap with ground truth is modestly acceptable, the model struggle with certain boundary regions or extreme segmentation errors. The high HD95 indicates that parts of the segmentation are much farther from the true boundary, undermining their overall reliability in precise delineation tasks. Moving on to the modified U-Net variants, both version 1 model and version 2 model achieve Dice scores close to the Native U-Net 0.8934 and 0.8984 respectively and retain low HD95 values near 1.54. These results imply that iterative architectural refinements in version 1 model and version 2 model maintain a robust performance on par with the baseline U-Net which potentially improving boundary alignment given the slightly higher Boundary Dice around which are 0.903 for version 1 and 0.908 for version 2. Finally, the proposed modified U-Net version 3 stands out by scoring the highest Mean Dice which is 0.9013 and maintaining a similarly high Boundary Dice which is 0.9082. Its HD95 of 1.52 is also marginally better than the native U-net's which indicates that it manages to combine high overlap accuracy with more precise boundary delineation. Consequently, this suggests that the additional architectural enhancements introduced in the proposed model led to subtle but meaningful performance gains relative to the other variants and the Cranext-based approaches.

TABLE I
QUANTITATIVE COMPARISON OF SKULL-IMPLANT ESTIMATION PERFORMANCE ACROSS BASELINE, LITERATURE, AND PROPOSED MODELS.

Model	Mean Dice	Mean Boundary Dice	Mean HD95
Native U-Net Model[16]	0.897296499	0.906607105	1.53531715
Cranext Pretrained Model[17]	0.877892674	0.877892674	8.362080358
Cranext Native Model[17]	0.874638764	0.874638764	8.404441038
Modified U-Net V1 Model	**0.893416421**	**0.902965986**	**1.540316383**
Modified U-Net V2 Model	**0.898378006**	**0.907950352**	**1.547305042**
Proposed Modified U-Net V3 Model	**0.901284902**	**0.908170876**	**1.519579183**

Figures 12 (a-c) show how effectively the proposed structure of version 3 model converts its numerical gains into clinically meaningful geometry. For three representative defects, a large temporo-parietal loss, a central vertex opening, and an elongated temporal-floor gap, the yellow digital meshes align perfectly with the surrounding cranial rims. When these meshes are 3-D printed in black color and seated on skull phantoms, the implants display virtually no gaps or step-offs, confirming that the model's lightened shallow encoder still captures the fine cortical contours required for a press-fit reconstruction. The printed surfaces remain smooth and faithfully follow the native calvarial curvature, where no stair-stepping or high-spot artefacts are visible which indicates that the decoder's channel-realignment preserves high-frequency detail during up-sampling. This fidelity is especially evident in the concave orbital roof and the convex parietal dome, shapes that had previously exposed weaknesses in the deeper, less balanced trial networks. Despite differences in defect size, location, and curvature, all three implants maintain edge-to-edge conformity and sit flush without manual adjustment, demonstrating that version 3 model reallocates capacity rather than diminishes it and produces files that are truly "print-ready," shortening the CAD workflow for patient-specific implants.

These visual findings correspond closely to the rankings in Table I. Version 3 model achieves the highest Mean Dice (0.9013) and Boundary Dice (0.9082) scores and the lowest HD95 (1.52 mm) of all tested models, metrics that explain the flawless margins seen in Figures 10 (a-c). By contrast, the other models, with high HD95 values, would manifest as conspicuous misalignments, precisely the boundary failures absent in version 3 model's demonstrations. Modified U-Net V1 and V2 approach the Dice performance of the native model but retain slightly higher HD95 values are ~1.54-1.55 mm; in practice, their prints occasionally required minor sanding to seat flush. Figure 11 reinforces this point where each black implant generated by version 3 model drops directly into its skull window without post-processing, whereas prints from the other models would need extensive manual reshaping. By delivering the best overlap accuracy and the tightest boundary tolerance, version 3 model not only validates the architectural decisions that shifted capacity from redundant shallow features to boundary-critical pathways but also demonstrates readiness for immediate surgical application.

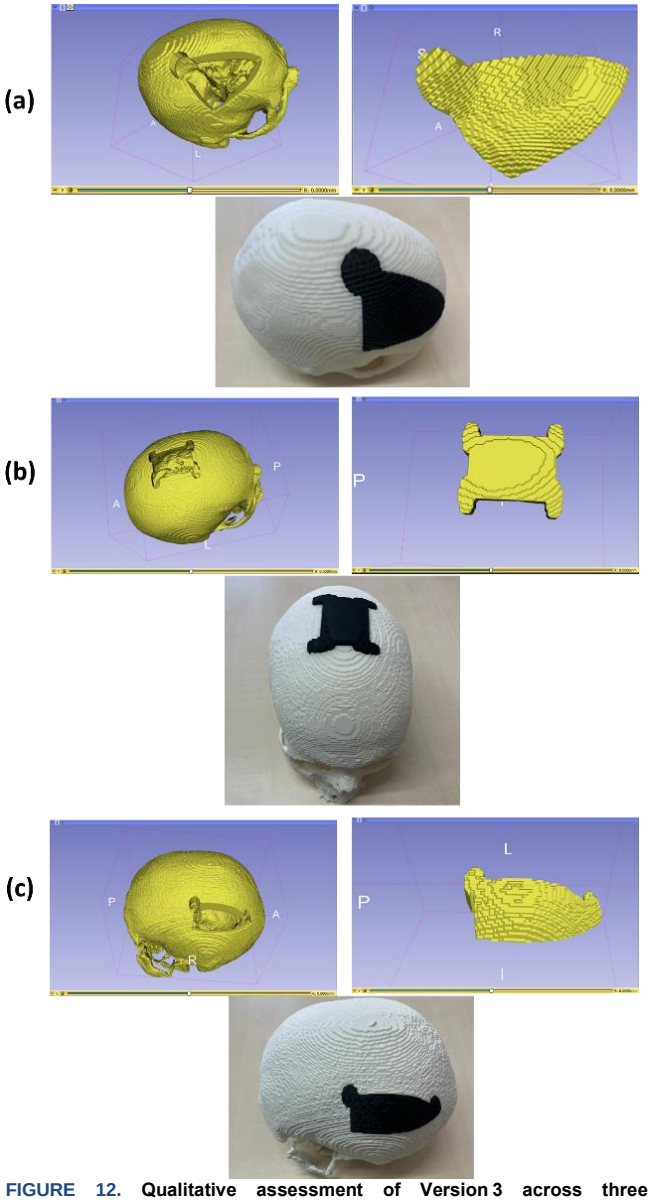

FIGURE 12. Qualitative assessment of Version 3 across three representative cranial defects: Yellow meshes are the network's digital predictions; black parts are the corresponding 3-D-printed implants seated on skull phantoms. From left to right: (a) large temporo-parietal defect, (b) box-shaped vertex craniectomy, and (c) elongated temporal-floor gap.

Figure 12 provides a tangible end-to-end verification of the proposed workflow by linking numerical segmentation metrics with physical fit quality. The panel depicts three representative skull phantoms, each harboring morphologically distinct cranial defects ranging from a smooth, convex frontal loss to an irregular temporal, parietal void. Adjacent to every phantom lies a matte-black implant fabricated directly from the version 3 network's STL output on an FDM printer with 0.2 mm nozzle and PLA test material. The implants are shown prior to seating so that their peripheral contours and internal lattice patterns remain visible. No intermediate CAD retouching, offset compensation, or post-print sanding was performed, thereby making the bench-top

scene a stringent test of "push-button" manufacturability. Where, the visual inspection of the printed part (see Figure 13) confirms that each implant mates flush with its corresponding defect border, leaving minimal daylight between polymer and bone phantom. This qualitative observation corroborates version 3's quantitative superiority in volumetric overlap and surface conformity reported in Section 5. Earlier modified U-Net variants frequently required manual rim thickening or edge smoothing to correct step artefacts and over-segmentation spikes, steps that added around 30 minutes of CAD labor per case and risked introducing user-dependent variability. In contrast, the implants in Figure 11 illustrate that version 3 model's predictions are already manufacturing-grade, meeting the ±1 mm clinical tolerance that craniofacial surgeons typically accept for intra-operative trimming.

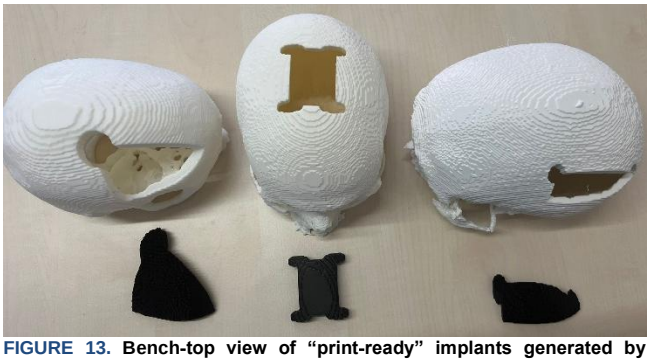

FIGURE 13. Bench-top view of "print-ready" implants generated by Version 3: Three skull phantoms with diverse defects are displayed alongside their black, 3-D-printed implants before seating. Each implant, produced directly from the network's STL output, matches the defect geometry without requiring manual CAD edits or post-print sanding, an outcome consistent with version 3's superior overlap and boundary metrics and unattainable with earlier Modified U-Net variants or other models

This study yielded three overarching insights. First, data quality is paramount. Public cranial-implant datasets are small and heterogeneous, limiting a network's ability to learn defect morphology. Conventional augmentations such as flip, rotate, and noise, inflate sample count but do not create new anatomical diversity.

Second, workflow integration matters as much as network design. Version 3 model's success hinged on a rigorously standardized pre-processing routine and a feedback loop with clinical specialists. Cropping, isotropic resampling, and fixed-size padding produced stable tensors for training, while surgeon input ensured that algorithmic improvements translated into implants that seat flush without manual adjustment. These outcomes point out the need for robust data pipelines on the engineering side and continuous anatomical oversight, whether from clinicians or AI "virtual consultants", on the clinical side.

Third, targeted architecture search pays dividends. Systematically testing baseline, deeper, and CraNeXt style networks revealed that boundary fidelity, not overall depth, drove fit accuracy. Rebalancing channels in Version 3 preserved high-frequency edge details and delivered the highest Dice and lowest HD95 scores, eliminating intra-operative trial-and-error. Future model work will therefore explore attention blocks and residual paths aimed specifically at boundary refinement rather than blanket depth increases.

Going forward, we will (i) fully automate high-fidelity defect rendering to enlarge the training corpus, (ii) incorporate boundary-aware attention and residual modules to push sub-millimeter accuracy, and (iii) develop an AI clinical agent for real-time anatomical validation and regulatory traceability. Collectively, these steps are expected to deliver a robust "push-button" system for on-demand cranial-implant design and AM, shortening surgical lead times and improving patient outcomes worldwide.

VI. CONCLUSION

This study introduces a fully automated framework for patient-specific cranial implant generation that spans image acquisition, volumetric pre-processing, deep-learning-based defect reconstruction, and physical validation through AM. The core technical contribution is the version 3 CNN model which is a channel-rebalanced 3-D U-Net that preserves high-frequency boundary information while reducing parameter count in the shallow encoder. Trained on a modest and heterogeneously sourced dataset, version 3 achieved state-of-the-art segmentation accuracy (Dice ≈ 0.90, BD ≈ 0.91) and the tightest boundary tolerance (HD95 ≈ 1.5 mm) among all models evaluated. Crucially, these numerical gains translated into tangible clinical value: implants printed directly from the network's STL files mated flush with three morphologically diverse skull phantoms, eliminating the need for labor-intensive CAD clean-up and demonstrating true "push-button" manufacturability.

Several strategic insights emerged. First, rigorous pre-processing such as cropping, isotropic resampling, fixed-size padding, and intensity normalization, proved indispensable for stable training and cross-scanner generalization. Second, in this application domain, native training outperformed transfer learning; pretrained CraNeXt weights failed to adapt, whereas random initialization allowed the network to specialize on cranial geometry. Third, data scarcity remains the chief bottleneck; even limited synthetic-defect rendering enhanced generalization, emphasizing the promise of large-scale, automated defect simulation.

Future work will therefore pursue two parallel tracks. On the data side, we will fully automate high-fidelity defect rendering and explore generative models to amplify dataset diversity. On the algorithmic side, we will integrate attention mechanisms, residual pathways, and boundary-aware loss functions to push accuracy toward surgical tolerances below 1 mm. In parallel, an AI "clinical agent" will be developed to provide real-time anatomical validation and regulatory traceability. Collectively, these advances aim to deliver a robust, point-of-care system for cranial implant design and AM, shortening surgical lead times, lowering costs, and enabling equitable access to bespoke neuro-cranial reconstruction worldwide.

ACKNOWLEDGMENT

The authors would like to acknowledge the support provided by the KIT-Publication Fund of the Karlsruhe Institute of Technology. Also, the authors would like to thank Dr. Tobias Müller, Senior Researcher at Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, for his support in this work.

REFERENCES

[1] A. Nazir, D. Vervoort, and C. L. Reddy, "From the first mile to the last: Challenges of the global surgical supply chain," (in eng), *Am J Surg*, vol. 222, no. 4, pp. 709-711, Oct 2021, doi: 10.1016/j.amjsurg.2021.03.033.

[2] H. S. Kim, S. H. Cho, D. H. Moon, and C.-H. Kim, "Risk factors for templating mismatch of uncemented stems in bipolar hemiarthroplasty for femoral neck fracture," *Scientific Reports*, vol. 13, no. 1, p. 21083, 2023/11/30 2023, doi: 10.1038/s41598-023-48538-y.

[3] D. D. Costa, K. Vickery, A. F. Tipple, and H. Hu, "Providing Sterile Orthopedic Implants: Challenges Associated with Multiple Reprocessing of Orthopedic Surgical Trays," *Hygiene*, vol. 2, no. 1, pp. 63-71doi: 10.3390/hygiene2010005.

[4] K.-H. Shin, J.-H. Kim, and S.-B. Han, "Greater Risk of Periprosthetic Joint Infection Associated with Prolonged Operative Time in Primary Total Knee Arthroplasty: Meta-Analysis of 427,361 Patients," *Journal of Clinical Medicine*, vol. 13, no. 11, doi: 10.3390/jcm13113046.

[5] A. Monetta *et al.*, "Prolonged operative time significantly impacts on the incidence of complications in spinal surgery," *Journal of Orthopaedic Surgery and Research*, vol. 19, no. 1, p. 567, 2024/09/14 2024, doi: 10.1186/s13018-024-05066-3.

[6] R. K. Whittaker *et al.*, "Component size mismatch of metal on metal hip arthroplasty: an avoidable never event," (in eng), *J Arthroplasty*, vol. 29, no. 8, pp. 1629-34, Aug 2014, doi: 10.1016/j.arth.2014.03.008.

[7] Z. Ashfaq, H. Ahmed, A. Khan, and A. Mufti, "Outcome of Early Cranioplasty in Trephine Syndrome or Paradoxical Brain Herniation: A Case Report and Literature Review," (in eng), *Cureus*, vol. 17, no. 3, p. e80922, Mar 2025, doi: 10.7759/cureus.80922.

[8] J. Maricevich *et al.*, "Functional and aesthetic evaluation after cranial reconstruction with polymethyl methacrylate prostheses using low-cost 3D printing templates in patients with cranial defects secondary to decompressive craniectomies: A prospective study," (in eng), *Surg Neurol Int*, vol. 10, p. 1, 2019, doi: 10.4103/sni.sni_149_18.

[9] V. Kopačin *et al.*, "Personalized 3D-printed cranial implants for complex cranioplasty using open-source software," (in eng), *Surg Neurol Int*, vol. 15, p. 39, 2024, doi: 10.25259/sni_906_2023.

[10] G. Pisaneschi, "Numerical and experimental exploration of patient-specific cranial implants with 3D-printed PEKK via Arburg Plastic Freeforming," *Progress in Additive Manufacturing*, 2025/04/12 2025, doi: 10.1007/s40964-025-01098-1.

[11] X. Chen, L. Xu, X. Li, and J. Egger, "Computer-aided implant design for the restoration of cranial defects," *Scientific Reports*, vol. 7, no. 1, p. 4199, 2017/06/23 2017, doi: 10.1038/s41598-017-04454-6.

[12] F. Altaf, S. M. S. Islam, N. Akhtar, and N. K. Janjua, "Going Deep in Medical Image Analysis: Concepts, Methods, Challenges, and Future Directions," *IEEE Access*, vol. 7, pp. 99540-99572, 2019, doi: 10.1109/ACCESS.2019.2929365.

[13] C.-T. Wu, Y.-H. Yang, and Y.-Z. Chang, "Three-dimensional deep learning to automatically generate cranial implant geometry," *Scientific Reports*, vol. 12, no. 1, p. 2683, 2022/02/17 2022, doi: 10.1038/s41598-022-06606-9.

[14] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, 2015: Springer, pp. 234-241.

[15] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, Cham, S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, Eds., 2016// 2016: Springer International Publishing, pp. 424-432.

[16] M. Wodzinski, M. Daniol, M. Socha, D. Hemmerling, M. Stanuch, and A. Skalski, "Deep learning-based framework for automatic cranial defect reconstruction and implant modeling," *Computer Methods and Programs in Biomedicine*, vol. 226, p. 107173, 2022/11/01/ 2022, doi: https://doi.org/10.1016/j.cmpb.2022.107173.

[17] T. Kesornsri *et al.*, "CraNeXt: Automatic Reconstruction of Skull Implants With Skull Categorization Technique," *IEEE Access*, vol. 12, pp. 84907-84922, 2024, doi: 10.1109/ACCESS.2024.3415173.

[18] M. Wodzinski, M. Daniol, and D. Hemmerling, "Automatic skull reconstruction by deep learnable symmetry enforcement," *Computer Methods and Programs in Biomedicine*, p. 108670, 2025/02/20/ 2025, doi: https://doi.org/10.1016/j.cmpb.2025.108670.

[19] L. Alzubaidi *et al.*, "A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications," *Journal of Big Data*, vol. 10, no. 1, p. 46, 2023/04/14 2023, doi: 10.1186/s40537-023-00727-2.

[20] G. Kaissis *et al.*, "End-to-end privacy preserving deep learning on multi-institutional medical imaging," *Nature Machine Intelligence*, vol. 3, no. 6, pp. 473-484, 2021/06/01 2021, doi: 10.1038/s42256-021-00337-8.

[21] R. M. Cichy, A. Khosla, D. Pantazis, A. Torralba, and A. Oliva, "Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence," *Scientific Reports*, vol. 6, no. 1, p. 27755, 2016/06/10 2016, doi: 10.1038/srep27755.

[22] M. Omar *et al.*, "The role of deep learning in diagnostic imaging of spondyloarthropathies: a systematic review," *European Radiology*, vol. 35, no. 6, pp. 3661-3672, 2025/06/01 2025, doi: 10.1007/s00330-024-11261-x.

[23] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60-88, 2017/12/01/ 2017, doi: https://doi.org/10.1016/j.media.2017.07.005.

[24] W. Yao, J. Bai, W. Liao, Y. Chen, M. Liu, and Y. Xie, "From CNN to Transformer: A Review of Medical Image Segmentation Models," *Journal of Imaging Informatics in Medicine*, vol. 37, no. 4, pp. 1529-1547, 2024/08/01 2024, doi: 10.1007/s10278-024-00981-7.

[25] M. Lackner, B. Puladi, J. Kleesiek, J. Egger, and J. Li, "A Semi-automatic Cranial Implant Design Tool Based on Rigid ICP Template Alignment and Voxel Space Reconstruction," *arXiv preprint arXiv:2404.15287*, 2024.

[26] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into Imaging*, vol. 9, no. 4, pp. 611-629, 2018/08/01 2018, doi: 10.1007/s13244-018-0639-9.

[27] M. D. Zeiler and R. Fergus, "Visualizing and Understanding Convolutional Networks," in *Computer Vision – ECCV 2014*, Cham, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., 2014// 2014: Springer International Publishing, pp. 818-833.

[28] C. T. Wu, Y. H. Yang, and Y. Z. Chang, "Creating high-resolution 3D cranial implant geometry using deep learning techniques," (in eng), *Front Bioeng Biotechnol*, vol. 11, p. 1297933, 2023, doi: 10.3389/fbioe.2023.1297933.

[29] Y.-L. Liao *et al.*, "Three-dimensional reconstruction of cranial defect using active contour model and image registration," *Medical & Biological Engineering & Computing*, vol. 49, no. 2, pp. 203-211, 2011/02/01 2011, doi: 10.1007/s11517-010-0720-0.

[30] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," (in eng), *Med Image Anal*, vol. 42, pp. 60-88, Dec 2017, doi: 10.1016/j.media.2017.07.005.

[31] L. Rundo and C. Militello, "Image biomarkers and explainable AI: handcrafted features versus deep learned features," *European Radiology Experimental*, vol. 8, no. 1, p. 130, 2024/11/19 2024, doi: 10.1186/s41747-024-00529-y.

[32] J. Nalepa, M. Marcinkiewicz, and M. Kawulok, "Data Augmentation for Brain-Tumor Segmentation: A Review," (in eng), *Front Comput Neurosci*, vol. 13, p. 83, 2019, doi: 10.3389/fncom.2019.00083.

[33] D. G. Ellis and M. R. Aizenberg, "Deep Learning Using Augmentation via Registration: 1st Place Solution to the AutoImplant 2020 Challenge," in *Towards the Automatization of Cranial Implant Design in Cranioplasty*, Cham, J. Li and J. Egger, Eds., 2020// 2020: Springer International Publishing, pp. 47-55.

[34] W.-M. Kung, I. S. Tzeng, and M.-S. Lin, "Three-Dimensional CAD in Skull Reconstruction: A Narrative Review with Focus on Cranioplasty and Its Potential Relevance to Brain Sciences," *Applied Sciences*, vol. 10, no. 5, doi: 10.3390/app10051847.

[35] J. Y. Abdullah, A. M. Abdullah, L. P. Hueh, A. Husein, H. Hadi, and Z. A. Rajion, "Cranial Implant Design Applying Shape-Based Interpolation Method via Open-Source Software," *Applied Sciences*, vol. 11, no. 16, doi: 10.3390/app11167604.

[36] J. Li *et al.*, "Back to the Roots: Reconstructing Large and Complex Cranial Defects using an Image-based Statistical Shape Model," *Journal of Medical Systems*, vol. 48, no. 1, p. 55, 2024/05/23 2024, doi: 10.1007/s10916-024-02066-y.

[37] M. Wodzinski, M. Daniol, and D. Hemmerling, "Improving the Automatic Cranial Implant Design in Cranioplasty by Linking Different Datasets," in *Towards the Automatization of Cranial Implant Design in Cranioplasty II*, Cham, J. Li and J. Egger, Eds., 2021// 2021: Springer International Publishing, pp. 29-44.

[38] A. G. Howard *et al.*, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[39] J. Deng, W. Dong, R. Socher, L. J. Li, L. Kai, and F.-F. Li, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 20-25 June 2009 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

[40] A. F. Agarap, "Deep learning using rectified linear units (relu)," *arXiv preprint arXiv:1803.08375*, 2018.

[41] S. Alam, N. K. Tomar, A. Thakur, D. Jha, and A. Rauniyar, *Automatic Polyp Segmentation using U-Net-ResNet50*. 2020.

[42] N. Anaraki, M. Tahmasbi, and S. R. Kheradpisheh, *Detecting Cadastral Boundary from Satellite Images Using U-Net model*. 2025.

[43] O. Oktay *et al.*, "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.

[44] D. Jha *et al.*, "Resunet++: An advanced architecture for medical image segmentation," in *2019 IEEE international symposium on multimedia (ISM)*, 2019: IEEE, pp. 225-2255.

[45] G. Chetty, M. Yamin, and M. White, "A low resource 3D U-Net based deep learning model for medical image analysis," *International Journal of Information Technology*, vol. 14, no. 1, pp. 95-103, 2022/02/01 2022, doi: 10.1007/s41870-021-00850-4.

[46] J. G. Li, Christina; Pepe, Antonio; Morais, Ana; Alves, Victor; von Campe, Gord; et al. *Head CT collection for patient-specific craniofacial implant (PSI) design. figshare.*, doi: https://doi.org/10.6084/m9.figshare.12423872.v1.

[47] J. Li *et al.*, "Synthetic skull bone defects for automatic patient-specific craniofacial implant design," *Scientific Data*, vol. 8, no. 1, p. 36, 2021/01/29 2021, doi: 10.1038/s41597-021-00806-0.

[48] O. Kodym *et al.*, "SkullBreak / SkullFix – Dataset for automatic cranial implant design and a benchmark for volumetric shape learning tasks," *Data in Brief*, vol. 35, p. 106902, 2021/04/01/ 2021, doi: https://doi.org/10.1016/j.dib.2021.106902.

[49] J. L. a. J. Egger. *SkullFix - MICCAI AutoImplant 2020 Challenge Dataset*.

[50] A. Dosovitskiy, P. Fischer, J. T. Springenberg, M. Riedmiller, and T. Brox, "Discriminative Unsupervised Feature Learning with Exemplar Convolutional Neural Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 9, pp. 1734-1747, 2016, doi: 10.1109/TPAMI.2015.2496141.

[51] Z. Eaton-Rosen, F. J. S. Bragman, S. Ourselin, and M. J. Cardoso, "Improving Data Augmentation for Medical Image Segmentation," 2018.

[52] A. Oliveira, S. Pereira, and C. A. Silva, "Augmenting data when training a CNN for retinal vessel segmentation: How to warp?," in *2017 IEEE 5th Portuguese Meeting on Bioengineering (ENBENG)*, 16-18 Feb. 2017 2017, pp. 1-4, doi: 10.1109/ENBENG.2017.7889443.

[53] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240-1251, 2016, doi: 10.1109/TMI.2016.2538465.

MAHMOUD SALEM received his M.Sc. degree in Computer Engineering and Systems from Ain Shams University, Cairo, Egypt, in 2020, and his B.Eng. degree in Electronics and Communication Engineering from Thebes Higher Institute of Engineering, Cairo, Egypt, in 2014. He is currently a Machine Learning and Robotics Scientist in the Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology, Karlsruhe, Germany. Previously, He worked as Computer Vision Research Engineer at MSD and EL2Labs in Cairo, Egypt. He has more than ten years of experience in research and development (R&D). Earlier, He also worked as R&D and embedded-software engineer roles with various institutions in Egypt. His research interests include computer vision, deep learning, robot learning, collaborative robotics, and Industry 4.0/5.0. During the past four years, he has contributed to international consortia such as SMERF (Interreg EU Funding, Germany/EU), the HEAT project (ZIM Funding, Germany/Taiwan), and Wertstromkinematik (KIT Internal Funding, Germany). He has also led industrial projects in the United States, Malaysia, Egypt, the Gulf region, and Sudan. He has authored or co-authored more than fifteen peer-reviewed articles. His work has been recognized with the Best Research Paper Award at CARV 2023 (Bologna), the Best Student Research Paper Award at KES-SDM 2019 (Budapest), and the Best Paper Award at ACHI 2018 (Rome). As an undergraduate, he secured podium places in four national robotics contests.

OMAR WAEL received his B.Sc. degree in Computer Science from the Faculty of Science, Cairo University, Egypt, in 2019, graduating with First Class honors. He later pursued a Postgraduate Diploma in Artificial Intelligence at the Information Technology Institute (ITI). He is currently a Senior Data Scientist at Data Gear, where he leads the development of machine learning solutions for Anti-Money Laundering (AML) systems used by international banks. His work involves enhancing AML detection models, improving data collection pipelines, and implementing supervised and unsupervised ML techniques to strengthen financial security measures. In addition to his industry expertise, he is deeply engaged in AI research. He is a Research Assistant at The American University in Cairo (AUC), contributing to an innovative Generative AI project at the School of Business. Simultaneously, he works as a Machine Learning Researcher at iStreamLabs, where he focuses on

cutting-edge computer vision projects. Previously, he worked as a Software Engineer at Phantasm Solutions, developing payment integrations and web applications.

MOATAZ ATTALLAH is the incoming Dean of the School of Aeronautical, Automotive, Chemical and Materials Engineering (AACME) at Loughborough University and holds a chair in advanced materials processing at the University of Birmingham. He is a Professor of Advanced Materials Processing (Metallic Materials). He received his PhD in metallurgy and materials science from the University of Birmingham (2007), BSc (highest honours) and MSc degrees from the American University in Cairo (AUC) Egypt, in mechanical engineering, and materials/manufacturing engineering respectively. Following his PhD, Moataz worked as a research fellow at the University of Manchester Materials Science Centre from January 2007, prior to his appointment as a lecturer at the School of Metallurgy and Materials at the University of Birmingham in June 2010. His research over the past 20 years focuses on developing a metallurgical understanding of the material-process interaction in advanced manufacturing processes (additive manufacturing, powder processing, friction joining, and superplastic forming) of metallic materials, focusing on the process impact on the microstructure and structural integrity development. His research is conducted through research partnerships with the UK Atomic Energy Authority (UKAEA), GKN Aerospace, Rolls-Royce plc, MBDA, European Space Agency (ESA), ITP, Aero Engine Controls, Johnson-Matthey, BAE Systems, TWI, Safran group companies (Safran Power Units and Safran Landing Systems), Meggitt, Honda R&D, IHI Corporation, Prima Power (Italy), M&I Materials, Magnetic Shields Limited, and the Manufacturing Technology Centre (MTC). His current grant portfolio as a PI includes projects funded by Innovate-UK, EPSRC, DSTL, the UKAEA, as well as several industrial projects, with a career total grant income of £ 15M as a PI. He has published over 200 journal and conference papers and 3 book chapters (h-index: 47), in addition to being a co-inventor for 5 patents. He has given keynote and invited talks in conferences, universities, and research centres in the USA, Canada, Japan, Germany, China, Italy, Finland, Australia, Singapore, Brazil, Mexico, UAE, and Egypt. His awards include the UK Ministry of Defence and the French Delegation Generale pour l'Armement Award in 2013, and the Safran Group Innovation Prize in 2014.

AHMED ELKASEER He is Full Professor of Manufacturing Engineering in the Department of Mechanical Engineering at the British University in Egypt (BUE). Ahmed Elkaseer earned his PhD in micro-milling from the Cardiff School of Engineering, Cardiff University, UK (2011), after graduating with honours in Production Engineering from Helwan University, Egypt (2001). Prof. Elkaseer's career spans several leading research centres: he served as a Post-Doctoral Researcher at Cardiff University, a Marie-Curie Experienced Researcher at IK4-Tekniker, Spain, and later a Senior Research Fellow at the Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology (KIT), Germany, where he mainly focused on 3D functional ink-jet printing for Industry 4.0 applications. With more than 20 years' experience, his expertise covers experimental, modelling, simulation, and optimisation studies across mechanical, EDM, laser, additive, and smart-manufacturing processes at conventional, micro-, and nano-scales, with a strong recent emphasis on intelligent manufacturing and Industry 4.0. He has led or contributed to projects funded by the UK EPSRC, multiple European Commission programmes, the Helmholtz Association, the British Council, and Egypt's STDF. Author of 130+ peer-reviewed publications, Prof. Elkaseer consults for industry and serves as guest editor, associate editor, editorial-board member, and reviewer for numerous manufacturing and materials journals. He regularly evaluates international research proposals and has chaired scientific committees and programme tracks at several international conferences, earning multiple awards for his contributions to advanced and intelligent manufacturing.