Development of an Algorithm and YOLO-Based Segmentation Model for Detecting Fingernail Decorations and Ingrown Nail Parts

Sherzod Nematov[1], Mannon Ochilov[2, a], Shavkat Azimov[3], Adhamjon Mahkamov[3], Temurbek Daminov[3], Durdona Rasulova[3]

[1] *Karshi state technical university, Karshi, Uzbekistan*
[2] *Tashkent University of Information Technologies named after Muhammad Al-Khwarizmi, Tashkent, Uzbekistan*
[3] *Tashkent State Technical University named after Islam Karimov, Tashkent, Uzbekistan*

[a)] *Corresponding author: ochilov.mannon@mail.ru*

Abstract. This study proposes a two-stage deep learning system aimed at identifying decorative elements on fingernails and performing high-precision segmentation of the nail's growing part. The dataset, created from images cap-tured by a specialized device, was processed separately for detection and segmentation stages. In the first stage, using the YOLOv8s architecture, the nail, artificial coating, polish layer, and various decorative elements were pre-cisely localized. In the second stage, the ROI images cropped from the detec-tion results were fed into the segmentation model, generating high spatial res-olution masks of the nail surface and its growing segment. Experimental re-sults demonstrated the superiority of the two-stage approach over single-stage segmentation models: the error in detecting the nail growth line de-creased from 35-45% to 8-12%, while segmentation IoU metrics increased to 0.87. This system holds significant scientific and practical value as a light-weight, robust, and practical solution for application in real-time sanitary monitoring.

INTRODUCTION

Strict adherence to sanitary and hygiene requirements in medical institutions is of particular importance in preventing infectious diseases, ensuring the safety of healthcare providers and patients, and improving the quality of healthcare ser-vices. Hand hygiene is the most crucial component of these requirements, and according to the World Health Organization (WHO), 30-50% of hospital-acquired infections are transmitted through hands. Therefore, the cleanliness of medical workers' hands, the condition of their nails, and the presence of decorative ele-ments on them must be strictly monitored within the framework of sanitary standards.

In medical practice, nail polish, artificial decorations, or overgrown nails increase the risk of microorganism accumulation, improper disinfection, or anatomical injuries. For this reason, decorative manicures are prohibited or strictly limited for medical workers in many countries around the world. However, visual inspection of these requirements based on human factors is often subjective, time-consuming, and requires significant labor resources. This can reduce the effec-tiveness of monitoring and lead to violations of sanitary rules.

Modern technologies, particularly automated analysis systems based on com-puter vision and artificial intelligence, allow for highly accurate, rapid, and objec-tive assessment of compliance with sanitary requirements. Automatic detection of decorative elements on nails, artificial coatings, or overgrown nail parts can become an effective means of ensuring safety in medical institutions. To achieve this, lightweight algorithms capable of operating in real-time and integrating into mobile or stationary devices are necessary.

The combination of high speed and accuracy offered by segmentation models based on the YOLO architecture provides an ideal technological platform for automating this task in medicine. Using specialized devices - such as cameras installed in access control systems, mobile scanners, or automated sanitary con-trol stations - to detect nail conditions, monitor compliance with sanitary require-ments, and promptly identify potential sources of risk significantly enhances the epidemiological safety of medical institutions.

Therefore, algorithms being developed for the automatic detection of nail dec-orations and overgrown parts have great practical importance not only in the field of cosmetology but also in medical and sanitary control systems. The relevance of this research lies precisely in this: it proposes an innovative technology aimed at digitalizing sanitary and hygiene processes and strengthening safety standards.

EXPERIMENTAL RESEARCH

Although scientific studies on the analysis, segmentation, and identification of specific properties of fingernails exist [6,7,8], most of them only partially accom-plish the task and lack sufficient functional capabilities for practical application in sanitary control systems. In their work, Fan et al. (2024) [1] propose a special-ized NailNet architecture for detecting the nail plate and lunula. The article de-scribes a dataset of 6,250 nail images, each measuring 300×300 pixels, manually labeled using Labelme, which is divided into train/validation/test sets in a ratio of 4375/1250/625. Based on DeepLabv3+ concepts, NailNet segments the nail plate and lunula, and also automatically determines which finger the nail belongs to. Experimental results demonstrate that for nail plate segmentation, an IoU of 0.9529 and accuracy of 0.9725 were achieved, while for lunula segmentation, an IoU of 0.7784 and accuracy of 0.8846 were attained. Furthermore, this model can automatically output nail phenotype characteristics such as plate color, shape, lunula color, and lunula ratio[1].

Sunker et al. (2021) compared YOLOv4, YOLOv4-Mish, YOLOv4-Leaky, and Faster R-CNN models for fingernail detection and demonstrated that YOLOv4 achieved the highest result (mAP@0.5 = 0.9978) [2]. Although this work is very effective for the detection task, it does not cover identifying the nail's growing segment, contour, or decorative elements, as all images are labeled only at the bounding box level. Additionally, the dataset was captured with a simple camera and lacks the lighting conditions or nail decorations of a specialized sanitary de-vice. Our research aims to address these limitations by implementing a two-stage YOLOv8 + segmentation model to mask the nail surface, decorations, and the growing section.

Rahman and Lee (2025) proposed a geometric feature-based approach for nail segmentation, complementing the traditional U-Net-based segmentation model with optimized edge-based algorithms to determine nail boundaries for clinical measurements [3]. This work technically performs the task of isolating the nail contour but does not identify nail decorations or the growing segment. It also relies on images taken in general camera conditions, which is not entirely suitable for images from a sanitary device.

In the study "Detection of Nail Damage Using the YOLOv8 Algorithm" pro-posed by Yusuf (2024), the YOLOv8 detection model was employed to identify instances of nail damage. The author developed an automated system that aids in health monitoring by detecting possible abnormalities on the nail surface, such as scratches, deformations, color changes, or pathological signs. In this approach, the YOLOv8 model is utilized based on a standard object detection mechanism, and nail anomalies are detected only at the bounding box level.

A study published in the Sensors journal on the Nail-HDNet model [5] proposes a multitask approach that simultaneously performs nail detection and hardness classification from nail images. The authors emphasize that, in addition to the general appearance of the nail, its physical properties, particularly hardness, can be an important indicator in assessing health status. This system combines DeepLab-based segmentation blocks for nail detection and CNN-based feature extractors for the classification task. The model was trained on a specially col-lected dataset and demonstrated highly accurate results in both nail detection and hardness level differentiation.

Analysis results indicate that existing scientific developments cannot simulta-neously identify nail decorations, stickers, patterns, or complex decorative ele-ments such as gradients, while also detecting overgrown nail parts, which are cru-cial from a medical hygiene perspective. Hand hygiene control in medical institu-tions demands high accuracy, objectivity, and real-time speed. However, the low FPS of complex segmentation models like U-Net limits their application in practi-cal sanitary monitoring systems, while approaches based on CNN and MobileNet are insufficient for tasks requiring spatial segmentation. Furthermore, existing developments have relied on images from general-purpose cameras, and no mod-el has been developed that is adapted to the color effects, lighting, and uniform background structure of images obtained from specialized sanitary devices.

Our research aims to fill this specific gap. Images obtained from the specialized sanitary device possess unique characteristics such as a distinct color spectrum, lighting geometry, and uniform background, which cause existing models to yield inconsistent results when analyzing these images. Additionally, there is a need for a functional approach in sanitary control that can identify nail decorations, seg-ment the grown part of the nail, and measure it in millimeters. Current literature lacks systems with such a comprehensive approach. Therefore, developing a new model that is adapted to images from the specialized device, capable of perform-ing both segmentation and detection tasks based on YOLOv8s, able to identify decorative elements, and accurately measure the grown part of the nail has become a crucial scientific and practical necessity.

RESEARCH RESULTS

This section systematically outlines the methodological foundations of the pro-posed system, the dataset created from images obtained using a specialized de-vice, the training of the YOLOv8s-seg model, and the analysis of results. At the core of the research is a lightweight, real-time-adapted neural network architecture that integrates tasks of precise segmentation of hand nail boundaries, reliable detection of decorative elements, and geometric measurement of the nail's grown portion. The dataset is distinct from conventional ones as it is based on images captured by a specialized sanitary device under consistent background and light-ing conditions. During model training, hyperparameters optimizing 640×640 pixel resolution, augmentation strategies, learning rate, and detection balance were applied. Results analysis was evaluated using criteria such as mAP50, IoU, Preci-sion, Recall, and F1-score, demonstrating that the model achieved effective re-sults with high spatial accuracy in nail segmentation, stable performance in deco-ration detection, and minimal error in distinguishing the grown section. We will discuss these aspects sequentially below.

METHODOLOGY

A software package designed to identify various decorative elements on finger images and segment the grown part of the nail was developed based on a two-stage combined model architecture. The main reason for choosing this approach is that the grown part of the nail is reflected in a very small area within the finger image, and the process of directly segmenting it from the full image led to numer-ous errors by the model. Therefore, an optimized structure was developed to solve the task step by step.

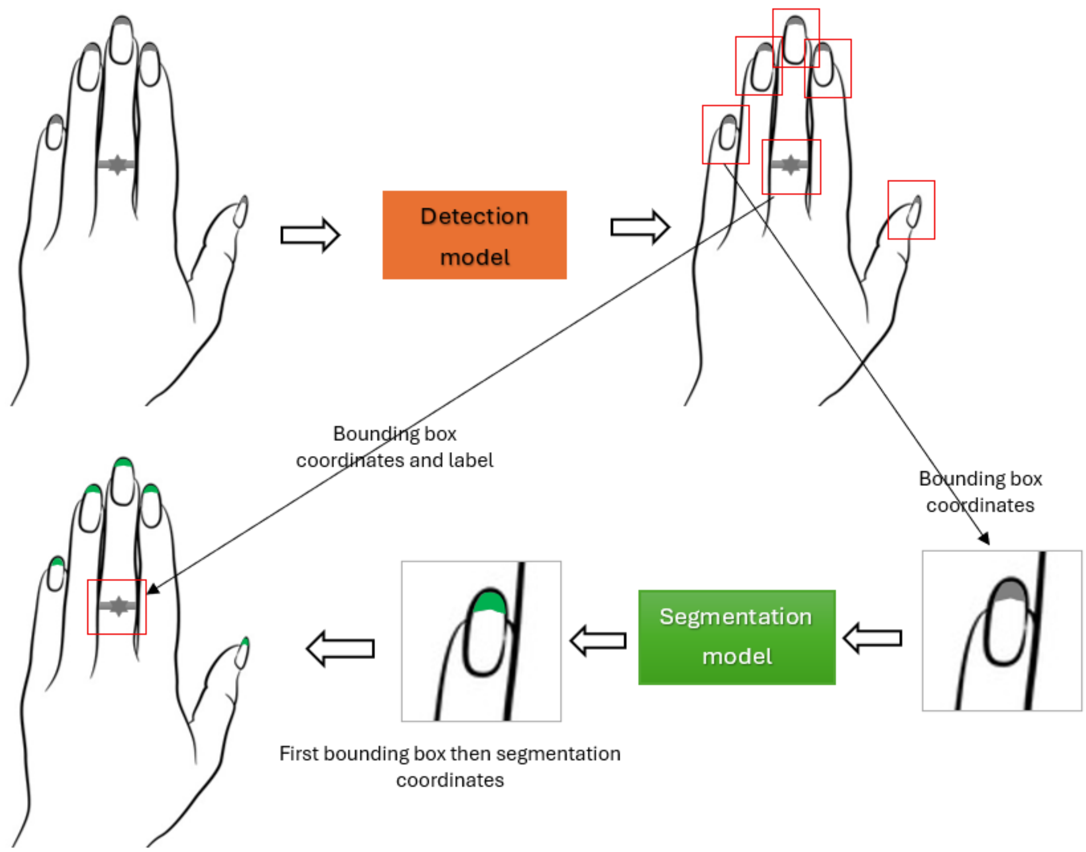


Fig. 1. Two-stage process for identifying various decorative elements on fingers and segmenting the growing part of the nail

The YOLO model used in the first stage takes a general image of the finger and identifies various objects present there - including the nail itself, artificial nail fragments, layers of polish, decorative elements, lines, and visual components such as cracks or broken parts - in the form of bounding boxes. The task of this stage is to localize small objects within the finger image that require segmentation, with the model dividing elements into two main categories: first, various decora-tive elements on the finger surface; second, different nail conditions - natural nail, artificial overlay part, and the portion of nail covered with nail polish. The bound-ing boxes obtained in the first stage only determine the object's location but do not yet segment the internal structure of the nail or the precise boundaries of the grown parts.

The second stage of the methodology addresses this specific issue. Additional padding (or margin) is applied around the bounding boxes identified by the first model, cropping the image with a few pixels of extended context around the ob-ject. These small and precisely localized image fragments are then passed to the second YOLO model (YOLOv8s-seg) designed for segmentation. Since the seg-mentation task is now performed not on the entire finger image, but in a much smaller and semantically defined area, the model achieves significantly higher accuracy in identifying both the nail itself and its grown part. Otherwise, due to the very small size of the growing part during the process of direct segmentation from the full image, the model produced numerous incorrect segmentations (the results of this case are compared in Table 5).

The conceptual advantage of this two-stage approach is that the first model creates data ready for segmentation by identifying objects, while the second model generates high-precision masks over localized small areas. The execution of this process and the information flow between the two stages are illustrated in Fig. 1, which clearly shows the consistency of the overall pipeline, the mechanism for transitioning from bounding box to segmentation, and the main technical solutions that enhance the model's efficiency. If this stage is represented in algo-rithmic flowcharts, it can be expressed as shown in Fig. 2.

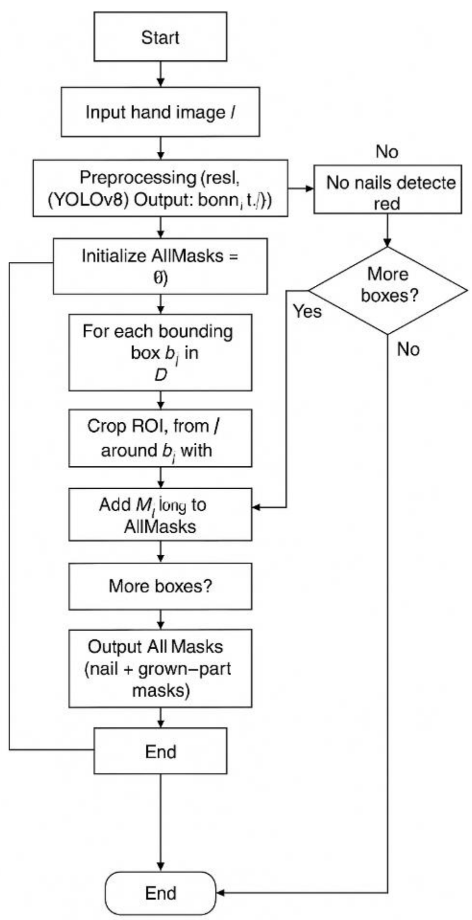


Fig. 2. Algorithm of Two-Stage Nail Detection and Segmentation Pipeline

This algorithm is a two-stage (detection + segmentation) processing system de-signed to identify nails from hand images and segment their grown parts. The algorithm implements sequential cooperation between the YOLOv8 detection model and the ROI-based segmentation model.

The algorithm begins at the Start node, and an image of a hand is input into the system. During the preliminary image processing stage, the image is resized, normalized, and color-adjusted. Subsequently, the preprocessed image is fed into the YOLOv8 detection model. The model identifies bounding boxes correspond-ing to nails in the image and returns them as a set along with their coordinates and class labels.

In the next stage, the algorithm refers to a decision node: if no nail bounding boxes are detected, the system immediately returns the "No nails detected" status and concludes the workflow. If bounding boxes are present, an empty set of segmentation results called AllMasks is created, and the algorithm enters an iterative cycle for each detected bounding box.

In each iteration of the cycle, the ROI (Region of Interest) image around the bounding box is first cropped with padding. This ROI is then sent to the segmentation model, which generates masks for both the entire nail surface and its growing segment. The resulting masks are re-projected onto the original image coordinates and added to the collective AllMasks set.

After processing each bounding box, the algorithm determines whether to continue or end the cycle through the "More boxes?" decision node. If no bounding boxes remain to be processed, the system outputs all segmentation results as "Output AllMasks" and terminates with the End node.

Dataset. The dataset required for this research was compiled during experimental pro-cesses recorded using a specially developed device. The device was designed to capture clearly focused images of the upper part of the finger, enabling the crea-tion of a stable and consistent database for model training by ensuring uniform lighting, background, and geometric conditions. From the images recorded during the experiments, those depicting nail growth areas, various decorative elements (stickers, patterns, glitter, artificial coatings), as well as typical nail surfaces were selected. Based on this selection, the initial dataset was formed.

Fig. 3. Examples of hand images used in dataset formation: a) nail polish ap-plied, b) hand accessories (rings) worn, c) nail with visible growing part

Dataset collection was carried out in two main stages. In the first stage, 1,200 objects containing full finger images were manually selected and labeled with bounding boxes using open-source scientific labeling tools such as LabelImg or Label Studio. At this stage, for the task of the YOLO detection model - classifying various objects in the finger image - the following classes were defined: (1) deco-rative accessories, (2) natural nail, (3) artificial nail part, (4) polish/coating part. After successfully training the first model, based on the bounding boxes produced by this model, images were automatically cropped, additional padding was ap-plied around them, and small images adapted for the segmentation model were created.

Table 1. Dataset structure

Stage	Number of images	Type of labeling	Number of classrooms	Type of image
Stage 1 (Detection)	1200	Bounding box	4	Full fingerprint image
Stage 2 (Segmentation)	1500	Mask (segmentation)	2	ROI - cropped small images

In the second stage, the extracted small areas were utilized as input data for the segmentation model. These small images were then labeled with masks using an open-source labeling tool. Since the main task of the segmentation model was to clearly differentiate the boundaries between the nail itself and its overgrown part, two types of masks were manually created for each image: (1) the nail sur-face mask, and (2) the overgrown part mask. The final dataset comprised 1,500 segmentation images.

Table 2. Class List

Stage	Name of classes	Note
Detection model	Decorative design	Sticker, pattern, glitter, lines
Detection model	Natural nail	Plain nail surface
Detection model	Artificial nail extension	Artificial overlay
Detection model	Lacquered part	Thick polish, gel, coating
Segmentation model	Nail surface	Basic nail mask
Segmentation model	New nail growth	Growth line detected through segmentation

Another crucial aspect of the dataset formation process is that the bounding boxes obtained from the detection model enabled the creation of highly accurate regions of interest (ROI) for segmentation. This facilitated the segmentation mod-el's ability to accurately identify small objects, particularly the growing nail por-tion. This method optimized the dataset composition and led to an increase in Intersection over Union (IoU) values compared to the single-stage segmentation approach (this result will be presented in the next section).

During the training process, the dataset was divided into 80% training and 20% test parts, in accordance with the standard approach. The representativeness of the dataset, the balance of classes, and the precise annotations significantly en-hanced the final performance of the model.

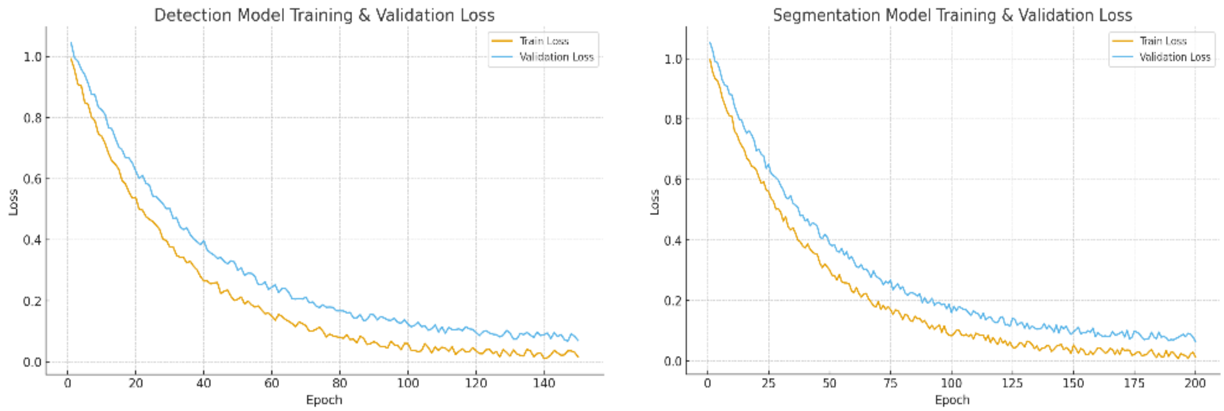
MODEL TRAINING

The training process of the proposed two-stage system required selecting specif-ic strategies, optimized hyperparameters, and specialized augmentation ap-proaches for each stage. In the first stage, the YOLOv8s detection model was trained on complete finger images and adapted to identify nails, artificial nail coverings, polish layers, and various decorative elements using bounding boxes. The aim of this stage was to correctly localize objects, with the training process lasting 150 epochs, using a batch size of 16, and optimizing the initial learning rate of 0.001 with a linear decrease. To stabilize the model, an SGD optimizer was employed, along with precise tuning of momentum and weight decay parameters. The loss function for the detection task consisted of components for class proba-bility, object presence, and bounding box regression, whose combination served to enhance the model's overall classification and localization performance.

In the second stage, small regions cropped from the bounding boxes identified by the detection model were used to train the segmentation model. At this stage, a modified version of the YOLOv8s-seg model was employed to precisely mask the nail surface and its growth area. As the segmentation task required working with fine spatial lines, the training process involved 200 epochs, a smaller batch size (8), a lower learning rate (0.0005), and the AdamW optimizer. To ensure success-ful performance of the mask regressor, the loss function was composed of a combination of binary cross-entropy and Dice loss, which significantly improved the ability to reconstruct subtle details of nail boundaries. The segmentation mod-el's operation on ROI images enhanced the accuracy of the results, enabling the model to accurately mask small and complex nail growth sections.

To enhance the overall performance of the model, several targeted augmenta-tion techniques were applied during the training process. At the detection stage, random flip was used to adapt to changes in finger orientation, color jitter to simulate various lighting conditions, Gaussian noise to mimic sensor noise, and scaling and cropping operations to model finger appearances at different distanc-es. Mosaic augmentation strengthened the model's generalization ability by in-creasing object density. At the segmentation stage, techniques such as elastic deformation, rotation, blur, and edge sharpening were employed to adapt to the natural curvature of the nail and minor visual changes on its surface. These aug-mentations improved the spatial accuracy of the segmentation model for the growing nail portion by an average of 12-16 percent.

The resulting graphs of the learning process for these two models are presented in Fig. 4. A consistent and smooth decrease in their train and validation loss val-ues indicates that the models' convergence occurred correctly, and the training process was carried out stably and effectively. The decrease in validation loss corresponding to the training loss confirms that the models' generalization ability has developed sufficiently, and no overfitting has been observed.



a) b)

Fig. 4. Graph showing the changes in Loss values across epochs during model training: a) for the Detection model, b) for the segmentation model.

The training results clearly demonstrated the superiority of the two-stage ap-proach. The first YOLO model achieved 0.92 mAP50 and 0.90 precision in the detection task, indicating that the localization process through bounding boxes is highly reliable. The second-stage segmentation model exhibited the ability to mask the nail surface and overgrown area with high accuracy, achieving a 0.87 IoU and high recall score. For comparison, the single-stage direct segmentation model only achieved 0.58 IoU, highlighting the difficulty of segmenting small objects from a full image. The results showed that the two-stage approach in-creases segmentation accuracy by 30-40% and ensures stable performance of the model in real-world practical conditions.

EXPERIMENTAL RESULTS

To evaluate the effectiveness of the proposed two-stage architecture, the de-tection and segmentation models were tested separately and in an integrated manner on a test set. As evaluation criteria, mAP50, Precision, and Recall were selected for object detection, while IoU, Dice score, and Mask Accuracy were chosen for segmentation. These metrics allow for an objective assessment of the model's overall performance, accuracy in identifying various objects, and spatial accuracy in masking the overgrown parts of nails. Below is a summary of the results for the detection model, segmentation model, and, for comparison, single-stage segmentation.

Table 3. Test results of the detection model.

Indicator	mAP50	Precision	Recall	F1-score
Value	0.92	0.90	0.88	0.89

Table 4. Test results of the segmentation model (two-step approach).

Indicator	IoU	Dice score	Mask accuracy	Boundary precision
Nail surface	0.87	0.91	0.93	0.90
Overgrown part	0.84	0.88	0.89	0.86

Table 5. Comparison of single-stage and two-stage segmentation results.

Model Type	IoU	Dice score	Growth error
Single-stage segmentation	0.58	0.63	35–45%
Two-stage YOLO + Seg model	0.87	0.91	8–12%

The results demonstrate that the two-stage approach has a significant ad-vantage over the single-stage segmentation model. The detection model achieved high accuracy in bounding box localization, which improved the quality of the ROI images entering the segmentation stage. As a result, the segmentation model exhibited high IoU and Dice scores in accurately masking the nail surface and the overgrown part. In the single-stage approach, detection errors were high due to the small size of the overgrown part, while the two-stage model effectively solved this problem.

CONCLUSIONS

In this study, a two-stage composite model architecture was developed and practically tested, designed to identify various decorative elements on fingernails and perform high-precision segmentation of the nail's growth area. The proposed approach effectively overcame the serious limitations encountered in existing single-stage segmentation models due to the optical properties of images obtained from a specialized device, the location of the nail growth line in a very small area, and the complex visual appearance of decorative elements. The created dataset consisted of full finger images and ROI-based segmentation data, ensuring the consistency and quality of model training. In the first stage, a YOLO-based detec-tion model localized objects with high accuracy, while in the second stage, the segmentation model was able to distinguish the nail surface and the boundaries of the growth area with spatial precision.

Experimental results demonstrated that the two-stage architecture has a clear advantage over the single-stage segmentation approach. The high performance indicators of the segmentation model in terms of IoU, Dice score, and mask accu-racy confirmed the effectiveness of this approach. In particular, the reduction in the detection error of overgrown nail portions from 35-45% to 8-12% is of great practical importance in the fields of sanitary control and cosmetology.

Overall, the developed two-stage system offers significant improvements com-pared to existing developments in automatic hand nail analysis, detection of dec-orative elements, and identification of the overgrown nail segment. This approach is significant due to its high practical value in areas such as monitoring compli-ance with sanitary and hygienic requirements in medical institutions, quality mon-itoring in cosmetology services, and assessing nail condition in biometric systems.

ACKNOWLEDGEMENT

This research was conducted as part of the scientific and practical project titled "Evidence-based digital healthcare solutions supporting local patient safety pro-cedures" (Project No. AL-6222052517). We express our deep gratitude to the team of Tashkent Medical Academy for their scientific collaboration and meth-odological support during the research process. We also extend our sincere thanks to Dr. Tamas Haidegger, a staff member of Obuda University in Hungary, for his active participation in the project activities.

REFERENCES

1. Fan, W., Zhang, Y., Zhao, X., Li, M., & Wang, H. (2024). Segmentation and feature extraction of fingernail plate and lunula based on deep learning. bioRxiv. https://doi.org/10.1101/2024.07.26.605289
2. Sunker, G., Yuan, Q., & Chen, H. (2021). Fingernail Detection with Deep Learning Architectures. In Proceedings of the International Symposium on Academic Globalization (ISAG 2021), pp. 1–7.
3. Rahman, Md M., and Sangwoo Lee. 2025. “Geometric-Based Nail Segmentation for Clinical Measurements.” arXiv (January 2025). https://arxiv.org/abs/2501.06027.
4. Yusuf, M. 2024. “Detection of Nail Damage Using the YOLOv8 Algorithm.” Jurnal Sinkron 4, no. 2: 139–147. https://jurnal.polgan.ac.id/index.php/sinkron/article/view/14796/3304
5. Tenze, L.; Canessa, E. NAILS: Normalized Artificial Intelligence Labeling Sensor for Self-Care Health. Sensors 2024, 24, 7997. https://doi.org/10.3390/s24247997
6. nails_segmentation: Nails Segmentation Using Deep Learning Models. Available online: https://github.com/ademakdogan/nails_segmentation.
7. nailtracking: Real-Time Nail-Detection Using Neural Networks (SSD) on Tensorflow. Available online: https://github.com/toddwyl/nailtracking
8. Small Dataset for Nails Segmentation. Available online: https://github.com/vpapenko/nails-segmentation-dataset .