İSTANBUL UNIVERSITY CERRAHPAŞA

ORIGINAL ARTICLE
Electrica 2024; 24(2): 425-435

An Image Processing-Based Approach for Reading Needle-Type Instruments on Aircraft

Fatma Gümüş, Can Eyüpoğlu

Department of Computer Engineering, Turkish Air Force Academy, National Defence University, İstanbul, Türkiye

Cite this article as: F. Gümüş and C. Eyüpoğlu, "An image processing-based approach for reading needle-type instruments on aircraft," *Electrica*, 24(2), 425-435, 2024.

ABSTRACT

For guaranteeing the safe and effective functioning of aircraft, image processing techniques can be a valuable tool to detect and evaluate aircraft panel values. In the pursuit of this objective, a dataset covering multiple aircraft models, various sessions, and different lighting conditions was compiled. Four tasks were examined through comparative analysis: object detection, display classification, needle masking, and needle angle detection. YOLOv8 demonstrated high performance in object detection and classification. In the classification task, the adaptability of needle-type device reading was examined by using the well-established models VGG16, Mobilenet V2, and Xception. Denoising autoencoder, U-net, and GrabCut methods were examined for needle masking, and the least squares method was applied to detect needle angle. As we move from the proof-of-concept phase to envisioning the development of an end-to-end system, this work provides significant analysis of image processing methodologies for reading aircraft dashboards.

***Index Terms*—**Aircraft analog indicators, cockpit dashboard reading, image processing, pointer needle detection.

I. INTRODUCTION

The Federal Aviation Administration (FAA) mandates the presence of recording devices on certain aircraft to collect crucial flight data, including the Cockpit Voice Recorder (CVR) and Flight Data Recorder (FDR) [1]. Exemptions exist for specific aircraft, but many countries, like Türkiye under Directorate General of Civil Aviation (DGCA) regulations [2], impose requirements for FDRs. The absence of FDRs on some planes complicates aviation safety investigations. To address this, lightweight recording equipment, such as cameras capturing instrument panels, offers a potential solution for obtaining critical information in incidents. This approach could significantly contribute to aviation safety, minimizing pilot distraction, automating maintenance processes, and enhancing anomaly prediction. Gumus and Eyupoglu [3] explored the motivation behind capturing and processing the cockpit dashboard extensively in their literature review.

The readings and measurements collected from the instruments and gauges on the control panel of an aircraft are known as aircraft panel values. These values are crucial for ensuring the safe and efficient operation of the aircraft. Typically, pilots or other crew members check these values during flight [3].

To monitor the performance of aircraft and identify any problems that may develop, image processing techniques can be used to detect and evaluate these panel values. Using machine learning (ML) techniques to analyze images of the control panel and identify indicator readings shown on the panel is one typical way to identify panel values with image processing. This is accomplished by training a ML model on a huge dataset of labeled images, where the labels represent the values shown on the panel. The model may then be utilized for classifying new control panel images and retrieving panel values. Another way is to analyze the images using computer vision methods to discover characteristics such as the forms and edges of the instruments and gauges on the panel. These characteristics may then be used to detect and extract panel values from images. There are inherent drawbacks in a dynamic flying environment, such as time-spatial oscillation during a video recording, although instrument reading in a constant environment is similar [3]. The aims and contributions of this study are as follows:

Corresponding author:
Can Eyüpoğlu

E-mail:
caneyupoglu@gmail.com

Received: December 2, 2023

Revision Requested: February 9, 2024

Last Revision Received: February 20, 2024

Accepted: March 26, 2024

Publication Date: May 13, 2024

DOI: 10.5152/electrica.2024.23190

- Meticulously curated a dataset featuring various aircraft models, sessions, and lighting conditions, providing a valuable and diverse resource for the research community in the domain of needle-type instrument reading.
- Proposed an image processing-based approach for reading needle-type instruments on aircraft.
- Explored the image processing algorithms' efficacy in the context of needle-type instrument reading by a task-based approach.

The remainder of the paper was organized as follows: Section II emphasized the need for recording and logging cockpit readings, listing the chosen literature organized by research motivation, and evaluating approaches utilized in various application areas, while Section III presented the used dataset and proposed approach. Section IV shows the experimental results and discussions. Finally, conclusions are summarized in Section V.

II. RELATED WORK

Various methods have been employed in recent studies for identifying and reading pointer instruments using machine vision. Salomon et al. [4] presented a dial meter reading approach incorporating YOLOv4 with an Xception-based AngReg model, achieving an identification rate of 98.9%. A dial meter reading system was developed by Devyatkin et al. [5] using the OpenCV library, involving illumination alignment, the Otsu method for image binarization, Hough transform, and angle processing. Ding and Zhang [6] utilized YOLOv3 for meter location, Hough transform for dial region extraction, and a combination of differentiable binarization and CRNN for scale extraction, achieving reading recognition through deflection angle calculation. Liu et al. [7] introduced a pointer meter reflection detection method based on YOLOv5s, perspective transformation, YUV color space, and enhanced k-means clustering with curve fitting, achieving a detection accuracy of 80.9%. Li et al. [8] focused on substation dial gauges, employing MSRCR for image enhancement, circle detection using the arc-support line segment based on the Hough transform, and RANSAC for pointer positioning and angle procedure.

Ma et al. [9] proposed an adaptive dual-pointer identification model incorporating Gaussian filtering, etching, dilation, Canny operator for edge detection, and the Hough transform for instrument dial detection. Chavan et al. [10] proposed an analog gauge reading method based on HSV color space segmentation, binary thresholding, morphological operations, Canny edge detection, Hough transform, and angle technique. In the study of Zhang et al. [11], a pointer meter reading recognition model was described, incorporating binarization, smooth denoising, edge detection, Hough transform for pointer line identification, silhouette approach for pointer position detection, and angle procedure stages, with a recognition accuracy rate of 98.07%. Li et al. [12] employed the Canny edge detector, Mask R-CNN, and angle method for SF6 pressure gauge reading. Sun et al. [13] utilized grayscale processing, Gaussian filtering, and the Otsu technique for image enhancement, extracting intersectant dual-pointer and pointer locations through the Hough transform. Hou et al. [14] presented a wireless sensor network-based pointer meter reading approach involving the Otsu method for image pre-processing, Hough transforms for round dial location, least squares method, and Zhang's thinning algorithm for pointer detection, and further steps for indicated value calculation.

Analyzing trends in measurements and natural phenomena, such as weather patterns, to detect indicators of potential natural catastrophes or aircraft irregularities is a valuable application. In the study of Nagarajan et al. [15], indicator values like altimeter, RPM gauge, fuel flow, and manifold pressure gauge were logged for future analysis. The study employed SSD Mobilenet v2 for dial identification and traditional image processing for dial reading, achieving 98.6% detection accuracy and 98% reading accuracy. Khan et al. [16] demonstrated this by monitoring airspeed and engine speed indications using two cameras, employing kernel techniques for simulation, and achieving an RMSE of 0 under real-world flight conditions.

Ensuring efficient flight operations and safety is crucial, with timely maintenance being a key factor. Addressing safety and maintenance, Hsiao et al. [17] applied the Canny edge detector and dynamic thresholding for dial edge detection, incorporating a "rotating needle mask" for needle angle determination. Zhang et al. [18] used an embedded image processing system with FPGA, employing grayscale transformation and template matching for dial reading. Similarly, Ricciardia and Minwalla [19] focused on airworthiness certification, utilizing a single camera aimed at the attitude indicator and torque gauge. They explored pixel intensity differences and binary thresholding for preprocessing and emphasized the importance of camera mounting location for effective area selection. Data extraction involved needle angle determination, aiming to reduce the time and cost of obtaining cockpit equipment permission. Extracting data, specifically determining the needle angle, play a crucial role in minimizing both the time and cost required to obtain permission for cockpit equipment. This process ensures a thorough check, comparing the values read from the system with those displayed on the dashboard screen to assess the device's reliability.

Pilot inattention was the number one cause of accidents according to several studies. It is important to monitor the devices in especially critical flight stages so that a warning can be issued in case of a risky maneuver. Pilot attention can shift when unexpected incidents occur, and the dashboard could be left unattended such as in the cases reported by NASA [20]. Moreover, the visual attention of pilots does not uniformly cease across all sections of the instrument panel [21]. Automatic monitoring of an aircraft's control panel during critical stages, such as takeoff and landing, is essential for ensuring safe flying performance [22-26]. Deviations from predicted parameter values require prompt pilot action, emphasizing the significance of monitoring flight parameter indicators. Tappan and Hempleman [27] monitored six indications using an array of cameras, relying on OpenCV as a toolbox. They successfully recognized needle locations within 2.9 degrees of predicted values, aiming to learn from and follow human pilot behavior. Tunca et al. [28] used YOLOv4 for real-time object detection of cockpit instruments, achieving an F1-score of 0.99 with GrabCut for needle mapping, aiming to reduce pilot effort through automated reading. Khan et al. [29] proposed a system for pilot assistance, using cameras to monitor airspeed and RPM gauges. They employed image subtraction and one-dimensional kernel convolutions for frame extraction, predicting needle angles, and categorizing readings using Bayesian theory.

This literature review including the studies mentioned above identified critical challenges and opportunities in needle-type instrument reading, specifically in the context of aircraft cockpit instrumentation. The unique challenges encompass factors such as sunlight, aerial maneuvers, and environmental conditions contributing to

color-contrast changes, reflections, and shadows. These intricacies highlighted the need for specialized methodologies in image processing tailored to the distinct conditions of cockpit image capture.

Moreover, the exploration unveiled the potential application of image processing in real-time settings, notably within Virtual Reality (VR) helmets. Ernst et al. [30] demonstrated the significant advantages of integrating essential cockpit indicators into a pilot's VR helmet during flight, emphasizing the potential for improved safety and operational efficiency.

Addressing the pervasive data scarcity challenge in the aviation domain, our review emphasized the difficulty in accessing confidential flight data recorded by instrument panel cameras. ML algorithms, while promising, face limitations in performance due to insufficient training data. Anticipating these challenges, we identified two key research directions for future investigations.

Firstly, the exploration of software simulations emerged as a viable solution, as evidenced by existing practices. Simulations offer a controlled environment for data acquisition, mitigating challenges related to data confidentiality and accessibility. Secondly, the growing trend of privacy-preserving ML, exemplified in applications like satellite image processing [31], presents a promising avenue for safeguarding sensitive information while harnessing the power of ML algorithms.

To bridge these gaps in the literature, our current study strategically addresses the identified challenges by leveraging a carefully curated dataset obtained through the GeoFS simulation environment. By adopting this approach, we aim to contribute valuable insights and methodologies to the field, paving the way for advancements in image processing techniques tailored to the unique demands of aircraft cockpit instrumentation. This strategic move aligns with the anticipated research directions, facilitating progress in overcoming data scarcity challenges and ensuring the confidentiality of sensitive flight data in future investigations.

III. MATERIALS AND METHODS

The materials utilized in our research, including the meticulously curated dataset, serve as the basis for our investigations. Additionally, the methods employed, ranging from image processing algorithms to transfer learning approaches, are detailed to provide a clear and transparent account of our experimental design.

A. Dataset Description

Our dataset construction, tailored for needle-type instrument reading in aircraft, adheres to essential criteria derived from our previous survey: 1) Accessibility: prioritizing ethical standards, privacy regulations, and open data sharing, our dataset encourages transparency and collaboration within the research community. 2) Aircraft diversity: to enhance generalizability, the dataset includes different aircraft types, capturing variations in cockpit configurations and facilitating a comprehensive evaluation of image processing algorithms. 3) Varied lighting settings: meticulously designed to replicate actual aircraft environments, the dataset incorporates diverse lighting conditions, ensuring our study addresses the challenges of needle-type instrument reading under varying illumination.

These requirements bolster the scientific rigor of our dataset, precisely meeting the demands of our research objectives and contributing to advancements in image processing methodologies for aviation contexts.

1) Data Collection

The dataset was collected using the GeoFS simulation environment [32]. GeoFS is an online flight simulation platform that replicates real-world flying conditions. It offers a global scenery with accurate terrain and weather representations. Users can choose from a variety of aircraft and navigate through different locations, experiencing changes in time of day, season, and geographical settings.

The inclusion of various aircraft types, including Cessna 172 Modern, Alphajet PAF, and Piper Cub, ensures a comprehensive evaluation of needle-type instrument reading across different cockpit configurations and flying characteristics. Sessions are conducted at different times of day (morning, afternoon, and evening) and seasons (Winter, Summer, and Spring), introducing variability in lighting conditions and weather patterns. This enables the assessment of algorithm performance under a spectrum of environmental scenarios. The data were collected in five sessions as indicated in Table I.

To enhance the usability of the dataset for our research objectives, we employed a meticulous annotation process. Skilled annotators carefully labeled each image to identify needle-type instruments. The annotation process involved selecting the area of interest (dashboard), identifying the dials (airspeed indicator, altimeter, attitude indicator, heading indicator, turn coordinator, vertical speed indicator, and other dials), and annotating the needle position of the airspeed indicator. We only focused on the needle position of the airspeed indicator out of our six named indicators because marking needle positions is a very labor-intensive task, and we wanted to deliver our proof-of-concept experiments swiftly.

2) Data Preprocessing

Before model training and evaluation, the dataset underwent preprocessing steps to ensure consistency and eliminate potential biases. We identified four tasks, each required specific preprocessing steps: Object Detection, Dial Classification, Needle Mask Composition, and Needle Angle Detection. The dataset splits for each task are presented in Table II.

For Object Detection, images were resized to 640 × 640 pixels, and pixel values were normalized to [0,1]. The dataset was split into training (70%), validation (20%), and test (10%) samples.

TABLE I. DATASET COLLECTION SESSIONS OVERVIEW

Session	Aircraft	Location	Time of Day	Season
1	Cessna 172 Modern	Chamonix—Alps—France	16:30	Winter
2	Cessna 172 Modern	Florence—Tuscany—Italy	09:30	Summer
3	Cessna 172 Modern	Sahara—Algeria	14:30	Spring
4	Alphajet PAF	Acropolis—Athens—Greece	12:00	Spring
5	Piper Cub	Acropolis—Athens—Greece	18:30	Spring

TABLE II. DATASET SPLITS FOR TRAIN, VALIDATION, AND TEST

Task	Train Samples	Validation Samples	Test Samples
Object Detection	5144	1470	734
Dial Classification	43 909	8782	2195
Needle Mask Composition and Needle Angle Detection	6953	Not applicable	456

TABLE III. DIAL CLASSIFICATION TASK CLASS DISTRIBUTION

Class	Number of Samples	Class Distribution
Airspeed indicator	6956	7%
Altimeter	7353	8%
Attitude indicator	6984	7%
Dashboard	6666	7%
Heading indicator	6961	7%
Other dial	48 411	50%
Turn coordinator	5898	6%
Vertical speed indicator	6713	7%

In Dial Classification, dials were cropped, resized to 88 × 88 pixels, and pixel values normalized to [−1,1]. An additional step for dial classification was data augmentation, which included random rotation and horizontal flipping. The dataset was then divided into training (80%), validation (16%), and test (4%) samples. Class labels and distribution for the task are as in Table III.

Needle Mask Composition and Needle Angle Detection shares dataset and preprocessing steps. For those tasks, dials were cropped, and the airspeed indicator was filtered. Further preprocessing involved resizing to 88 × 88 pixels, grayscale transformation, and Otsu's Thresholding [33]. The training set constitutes 94% of the dataset, while the test set comprises 6%, with no validation set specified for this task.

B. Breakdown of the Tasks

The study examined four tasks: Object Detection, Dial Classification, Needle Mask Composition, and Needle Angle Detection. In contrast to proposing an end-to-end system where tasks are interconnected sequentially, each task was independently examined. An end-to-end system typically involves a sequential flow, taking and processing output from the previous task, such as framed video input, object detection of dashboard and dials, dial classification of the detected objects, composing a needle mask, and finding needle angle. However, our approach deviates by investigating each task in isolation, assuming perfect input accuracy from the preceding task. This method allows for a focused analysis of each task's performance, contributing to a nuanced understanding of individual components and facilitating targeted improvement strategies.

Fig. 1 depicts a typical end-to-end needle-type instrument reading system. In this work, the four tasks were handled independently. We proposed to perform object detection and classification tasks in one go with the YOLO v8 model [34]. We also presented further analysis on the classification task separately, utilizing transfer learning on different models: VGG16 [35-37], MobileNet V2 [38], and Xception [39]. A visual comparative result analysis of needle masking was performed on a denoising autoencoder, U-net model [40], and GrabCut [41]. Finally, the least-square method was applied to detect the needle angle.

1) Models and Algorithms

Most object detection algorithms employ regions to locate objects within an image, focusing on areas with high probabilities of containing the object, a.k.a. two-stage detection. YOLO (You Only Look Once) is a one-stage detector. It utilizes a single convolutional network to predict both bounding boxes and class probabilities for these boxes, providing a unique approach compared to region-based algorithms. YOLOv8, being a leading-edge state-of-the-art model released in 2023, advances upon its predecessors by introducing new features and enhancements. It accommodates a wide array of vision AI tasks, such as detection, segmentation, pose estimation, tracking, and classification.

YOLO is a one-stage algorithm; it detects and classifies at the same time. Therefore, classification is done in its inherent architecture. For proof-of-concept inspection of the four identified tasks individually, we also performed transfer learning with some of the well-known image classification models (Table IV). The dials were extracted by the reference of the ground-truth bounding boxes, and a separate classification dataset was formed. Each base model was trained for 10 epochs, and the initial weights of Imagenet were transferred.

Denoising convolutional autoencoders were employed with the goal of removing noise (background) and enhancing the clarity of the needle on the instruments. Autoencoders consist of an encoder and a decoder. The encoder compresses the input image into a

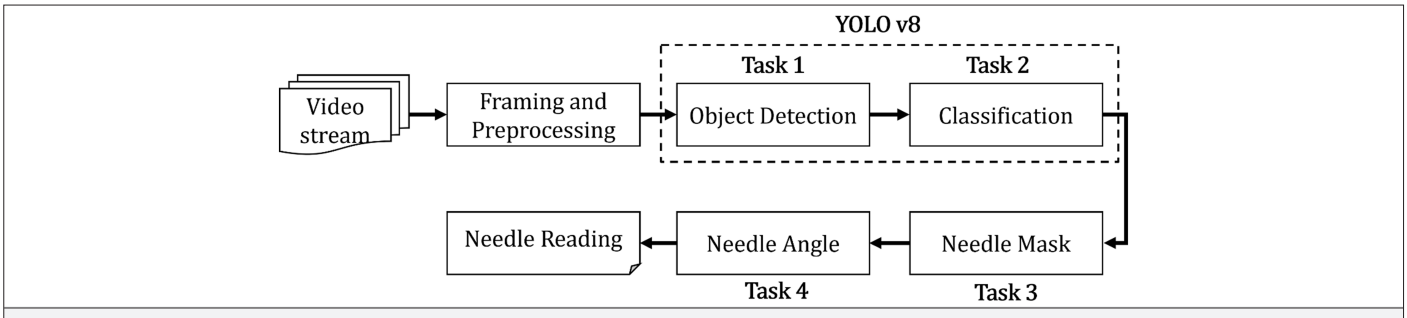

Fig. 1. A typical end-to-end needle-type instrument reading system.

TABLE IV. BASE MODEL INFORMATION FOR THE CLASSIFICATION TASK

Base Model	Model Size	Parameter	Depth
VGG16	528 MB	138.4 M	16
Mobilenet V2	14 MB	3.5 M	105
Xception	88 MB	22.9 M	81

latent space representation, and the decoder reconstructs the image from this representation while reducing noise. The autoencoder was composed of two convolutional layers (each 32 units) with max pooling in the encoder and two deconvolutional layers (each 32 units) in the decoder. It was designed for 2D input; thus, grayscale conversion was needed at the preprocessing step.

U-Net is designed for semantic segmentation tasks. Its architecture incorporates a contracting path (encoder) to capture context and a symmetric expanding path (decoder) for precise localization. U-net (Fig. 2) was composed of four downsampling blocks (64, 128, 256, and 512 units), a bottleneck (1024 units), and four upsampling blocks (512, 256, 128, and 64 units).

GrabCut is an interactive segmentation algorithm that separates an image into foreground and background based on user-specified initializations. Unlike the previous two segmentation algorithms, GrabCut requires a rough bounding box selection around the needle manually. This bounding box serves as an initial estimate of the foreground. Each pixel is considered a mixture of Gaussians, with parameters learned from the image data. The segmentation problem is formulated as an energy minimization task, where the energy function is defined based on color and spatial proximity.

GrabCut utilizes a Gaussian Mixture Model (GMM) to represent each pixel in an image, with each Gaussian component modeling distinct color distributions. Parameters are learned during initialization based on a user-provided bounding box. The GMM captures color variability, enabling the algorithm to handle complex foreground and background colors. The segmentation problem is framed as an energy minimization task, considering both color and spatial proximity. Through iterative updates, GrabCut refines the GMM parameters and segmentation based on the bounding box, enhancing accuracy for an interactive and precise foreground-background segmentation result.

The procedure in GrabCut is neither entirely manual nor automatic. The system operator can assign an area of interest to the algorithm based on the make and model. In our experiments, the optimal area of interest is identified around the base of the needle, achieved by dividing the device bounding box into a 3-by-3 grid and selecting the central part. Therefore, once the system operator establishes the ideal grid for the device under monitoring, the process becomes automatic for aircraft of the same make and model. Without selecting an initial interest area, the algorithm extracts a clutter of foreground information, not solely the needle. In such cases, significantly more post-processing would be required.

The final task detection of needle angle is a relatively easy task, presuming a correct needle mask got delivered from the pipeline. The least squares method for needle angle detection is a mathematical approach that aims to find the best-fitting line through a set of data points representing the position of a needle. In the context of needle angle detection, the method is applied to determine the orientation of the needle based on the positions of its tip and tail. As it was assumed that a suitable input is provided via the image processing pipeline, the tail should be in the center of the dials, and the other end of the tail.

2) Evaluation Metrics

Evaluation of object detection primarily entails bounding box loss (Equations 1 and 2), classification loss (Equation 3), and depth focal loss. The box loss measures the difference between the predicted bounding box and the ground truth coordinates. It is often computed using a regression loss, such as Mean Squared Error (MSE).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \check{y}_i\right)^2 \tag{1}$$

In calculating MSE for object bounding boxes, n is the number of bounding box coordinates (4 for x, y, width, and height in object

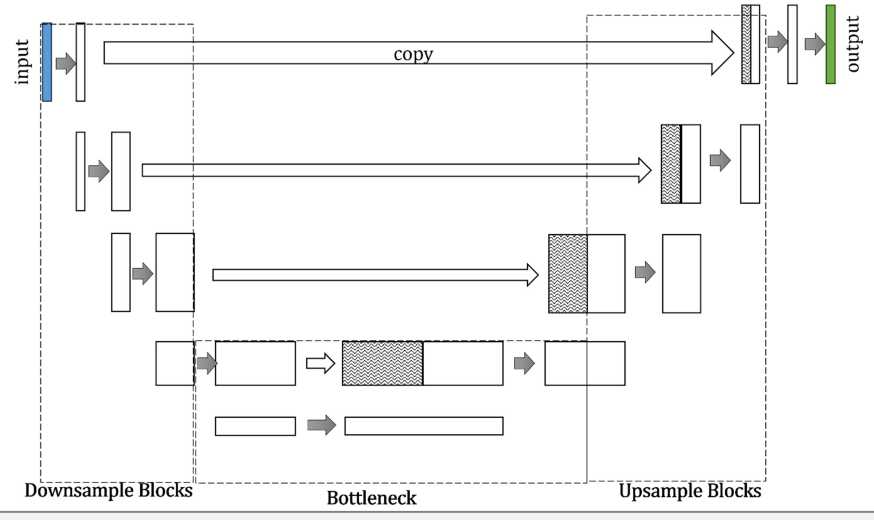

Fig. 2. The general architecture of U-net.

detection), where y_i is the ground truth bounding box coordinate and $\check{y}_i$ is the corresponding predicted bounding box coordinate. Bounding box loss is the final MSE value that represents the average squared difference between predicted and ground truth bounding box coordinates (Equation 2).

$$Box\,Loss = \frac{1}{4}\left(MSE_x + MSE_y + MSE_w + MSE_h\right) \tag{2}$$

The final box loss is the average MSE for the four ground-truth, and prediction values where $MSE_x = \left(\check{x} - x_{gt}\right)^2$, $MSE_y = \left(\check{y} - y_{gt}\right)^2$, $MSE_w = \left(\check{w} - w_{gt}\right)^2$, $MSE_x = \left(\check{h} - h_{gt}\right)^2$ as *gt* stands for "ground-truth", *hat* symbols for predictions, and *w* and *h* for width and height respectively.

The classification loss assesses the model's object classification performance, determined by computing the softmax cross-entropy loss. During the classification process, the model predicts class probabilities for each bounding box, generating a probability distribution across all possible classes for every identified object. This distribution is obtained through the application of a softmax activation.

$$Crossentropy\,Loss = -\sum_{i=1}^{C} y_i . \log\left(\check{p}_i\right) \tag{3}$$

Equation 3 describes a basic cross-entropy loss function where *C* is the number of classes, y_i is the ground truth probability for class *i*, and $\check{p}_i$ is the respective predicted probability.

Focal Loss is an enhanced version of Cross-Entropy Loss designed to handle class imbalance. It achieves this by assigning more weight to challenging or easily misclassified examples (e.g., background with noise, partial objects) and down-weighting simpler examples (e.g., background objects). The depth focal loss is specific to YOLOv8 and contributes to the model's ability to estimate distances; the exact equation was not released by the creators of the model as of the date.

In the context of object detection, mean Average Precision (mAP) is a widely used metric to evaluate the performance of models like YOLOv8. Two common variations of mAP are mAP50 and mAP50-95, which are calculated based on the Intersection over Union (IoU) metric.

The average precision for a specific class and IoU threshold is calculated by computing the precision–recall curve and taking the area under the curve (AUC). The precision-recall curve is created by varying the confidence threshold for object detection and calculating precision and recall values at each threshold. Precision and recall are calculated by Equations 4 and 5. The average precision is then obtained by interpolating the precision-recall curve and calculating the area under it.

$$Precision = \frac{True\,Positives}{True\,Positives + False\,Positives} \tag{4}$$

$$Recall = \frac{True\,Positives}{True\,Positives + False\,Negatives} \tag{5}$$

The two mAP variants are calculated by Equations 6 and 7, where *C* is the number of classes, AP_{50}^{c} is the average precision at IoU of 50 for class *c*, and AP_t^c is the Average Precision at IoU threshold *t* for class *c*. The sum is taken over the IoU threshold range from 50 to 95, and the division by 11 represents the number of thresholds in that range.

$$mAP_{50} = \frac{1}{C}\sum_{c=1}^{C} AP_{50}^{c} \tag{6}$$

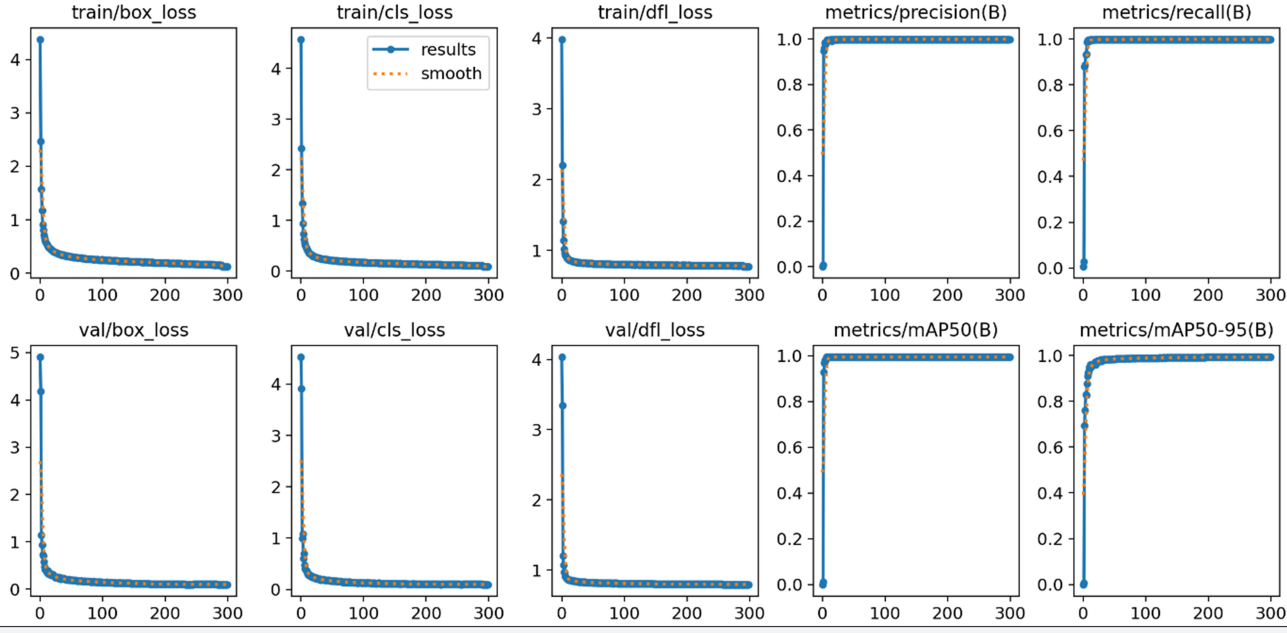

Fig. 3. YOLO v8 object detection training, where box_loss is bounding box loss, cls_loss is classification loss, dfl_loss is depth focal loss, mAP is the mean average precision metric at IoU of 50 (mAP50), and at the IoU range of 50–95 (mAP50-95).

Fig. 4. Dashboard illustration for reference.

$$mAP_{50-95} = \frac{1}{C}\sum_{c=1}^{C}\frac{1}{11}\sum_{t=50}^{95}AP_t^c \quad (7)$$

VI. EXPERIMENTAL RESULTS AND DISCUSSION

In our pursuit of dataset diversification, we meticulously curated data from five distinct sessions, involving three different aircraft models and varied lighting conditions. The YOLOv8 model, renowned for its robustness, demonstrated exceptional efficiency, achieving peak performance within a few epochs, as illustrated in Fig. 3. In the evaluation of the test set, YOLOv8 exhibited remarkable accuracy, boasting scores of mAP 99.5%, precision 99.9%, and recall 99.9%. Furthermore, its robust generalization was validated on additional held-out images from a brief sixth flight session, affirming consistent and accurate predictions. For clarity, an illustrative example, complete with ground-truth references, is presented in Figs. 4 and 5.

In the dial classification task, we leveraged transfer learning by employing three base models: VGG16, Mobilenet V2, and Xception. The accuracy results, as depicted in Table V, showcase the performance across three different scenarios:

- Without re-training: The base models were utilized without undergoing any additional training on the specific dial classification task. The models retained the knowledge gained from their pre-training on large-scale datasets but were not adapted to the nuances of the dial classification domain. Consequently, the performance in this setting reflects the models' generic understanding without specialization for the task at hand.
- Frozen re-training: The base models were incorporated into the dial classification task architecture, and certain layers were frozen, preventing further updates to their weights. This approach allowed the model to adapt to the dial classification task while retaining the valuable knowledge acquired during pre-training on broader datasets.
- Fine-tuning: This scenario involved training the entire model, including all layers, on the dial classification dataset. Unlike the 'Frozen Re-training' scenario, here, all parameters of the base model were open for updates.

As the results in Table V indicate, classification without re-training meant no adaptation to the dial classification domain. All models suffer, but especially Xception. Re-training the model on a frozen setting or not did not make a major difference.

Fig. 5. YOLOv8 prediction on a sample. All dials were correctly classified with the annotated confidence scores.

TABLE V. ACCURACY RESULTS FOR THE CLASSIFICATION TASK

Base Model	Without Re-training	Frozen Re-training	Finetuning
VGG16	0.1427	0.9945	1.0
Mobilenet V2	0.1113	0.9963	1.0
Xception	0.0333	0.9858	0.9995

Fig. 6. Input (left) and desired output (right) for needle mask task.

Needle masking was a particularly hard problem to tackle. We identified two main reasons:

- Limited contrast: needles may have limited contrast against the background or dial, especially in cases of poor lighting conditions or reflective surfaces. This lack of contrast makes it challenging for computer vision algorithms to accurately identify and mask the needles.
- Noise and interference: noise in images or interference from other elements on the dashboard, such as markings or shadows, can further complicate the task of accurately masking needles.

An example of input-desired output is shown in Fig. 6. We manually annotated the training set for the autencoder and U-net models. As shown examples in Fig. 7, the outputs were not ideal. Denoising convolutional autoencoder and U-net was unsuccessful. Training these models further, not only did not provide any improvement, but at 10 epochs the needle itself disappeared completely. In contrast, GrabCut was near perfect; however, it required manual annotation of the needle bounding box.

As our study centers on data collection and proof-of-concept analysis, we did not annotate the angles of the needles, precluding us from conducting an angle error analysis at this stage. It is noteworthy to mention, however, that the least squares method emerged as

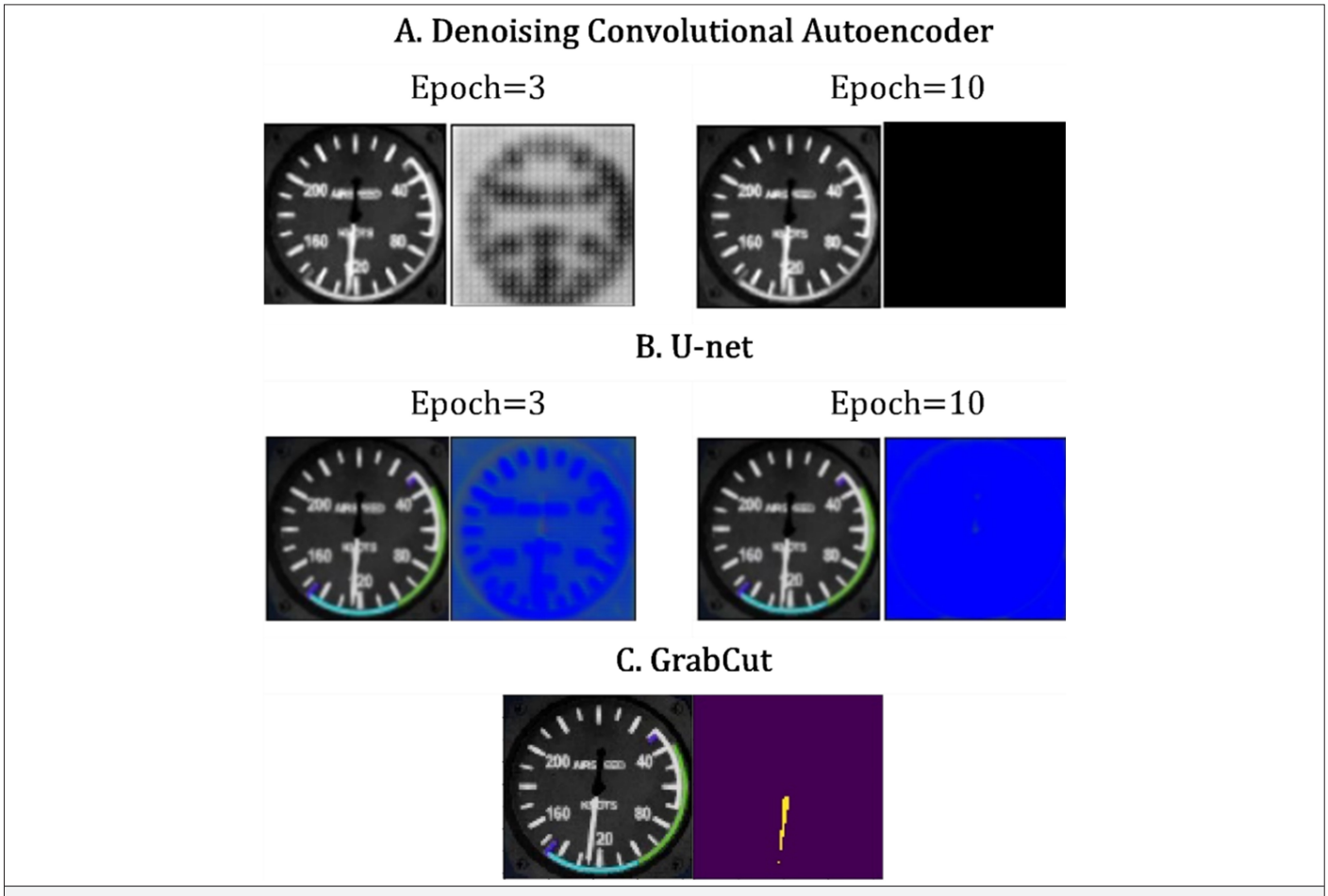

Fig. 7. Input and needle mask outputs: (a) Denoising convolutional autoencoder input and output at epochs 3 and 10. (b) U-net architecture input and output at epochs 3 and 10. (c) GrabCut algorithm input and output.

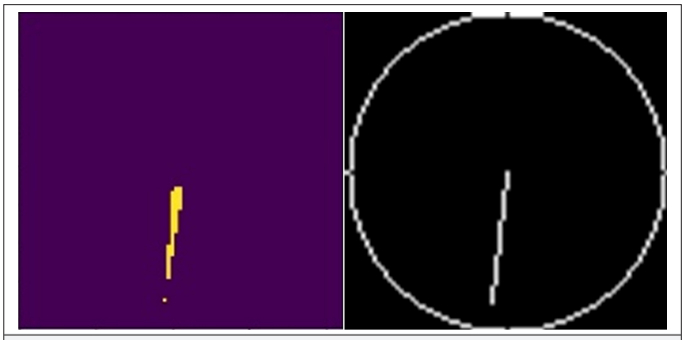

Fig. 8. Input (left) and calculated output (right) for the needle mask task, which was −96.70°.

an effective technique for angle estimation, leveraging the provided needle masks (Fig. 8).

V. CONCLUSION

In this study, we undertook a comprehensive exploration of needle-type instrument reading in aircraft through the collection and analysis of a curated dataset. Emphasizing the independent execution of key tasks such as object detection, dial classification, needle mask composition, and needle angle detection, we laid the foundation for a systematic understanding of image processing methodologies in the aviation context.

The dataset, comprising data from diverse sessions, aircraft models, and varying lighting conditions, serves as a valuable resource for future research endeavors. Our approach, rooted in a proof-of-concept framework, aimed to showcase the efficacy of individual tasks, highlighting the capabilities of advanced models like YOLOv8 in this specialized domain.

On the needle mask task, three approaches were examined separately. Denoising autoencoder and U-net did not yield desirable outcomes, and GrabCut required manual preprocessing. The three methods are to be used as a pipeline in a future study, with different classes of annotated dials. These algorithms complement each other: Denoising autoencoders enhance image quality, U-Net performs precise segmentation, and GrabCut refines the obtained masks for the most accurate representation of needle-type instruments. By experimenting with these approaches, the goal is to optimize the process of needle mask composition, ensuring high-quality masks that can be further utilized for needle angle detection and other relevant analyses in the context of aircraft instrument reading.

While our current focus remains on discrete task analysis, the prospect of developing an end-to-end system looms on the horizon. The seamless integration of individual tasks into a cohesive, interconnected system represents the logical next step, promising a holistic solution for needle-type instrument reading on aircraft dashboards.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – F.G., C.E.; Design – F.G., C.E.; Materials – F.G., C.E.; Data Collection and/or Processing – F.G., C.E.; Analysis and/or Interpretation – F.G., C.E.; Literature Search – F.G., C.E.; Writing – F.G., C.E.; Critical Review – F.G., C.E.

Declaration of Interests: The authors have no conflicts of interest to declare.

Funding: The authors declare that this study received no financial support.

REFERENCES

1. Cockpit Voice Recorders (CVR) and Flight Data Recorders (FDR). [Online]. Available: https://www.ntsb.gov/news/Pages/cvr_fdr.aspx [Accessed: March 20, 2023].
2. Directorate General of Civil Aviation, Homepage. [Online]. Available: https://web.shgm.gov.tr/en [Accessed: March 20, 2023].
3. F. Gumus, and C. Eyupoglu, "A survey of prominent image processing research on needle-type instrument reading on aircraft," 10th International Conference on Recent Advances in Air and Space Technologies (RAST), Istanbul, Türkiye, Vol. 2023, 2023. [CrossRef]
4. G. Salomon, R. Laroca, and D. Menotti, "Image-based Automatic Dial Meter Reading in unconstrained scenarios," *Measurement*, vol. 204, p. 112025, 2022. [CrossRef]
5. A. V. Devyatkin, A. R. Muzalevskiy, and A. S. Morozov, "Computer vision system for image-based automated dial meter reading," 2022 XXV International Conference on Soft Computing and Measurements (SCM), Saint Petersburg, Russian Federation, 2022. [CrossRef]
6. W. Ding, and B. Zhang, "Detection and recognition method for pointer-type meter based on deep learning," 2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology (ICCASIT), Changsha, China, 2021. [CrossRef]
7. D. Liu, C. Deng, H. Zhang, J. Li, and B. Shi, "Adaptive reflection detection and control strategy of Pointer Meters based on Yolov5s," *Sensors*, vol. 23, no. 5, p. 2562, 2023. [CrossRef]
8. D. Li, W. Li, X. Yu, Q. Gao, and Y. Song, "Automatic reading algorithm of substation dial gauges based on coordinate positioning," *Appl. Sci.*, vol. 11, no. 13, p. 6059, 2021. [CrossRef]
9. T. Ma, T. Li, M. Zhang, C. Zhang, and K. Shi, "An adaptive method for recognition of dual-pointer mechanical instrument based on Hough Transform," *IOP Conf. S. Earth Environ. Sci.*, vol. 632, no. 3, p. 032038, 2021. [CrossRef]
10. S. Chavan, X. Yu, and J. Saniie, "High precision analog gauge reader using optical flow and computer vision," 2022 IEEE International Conference on Electro Information Technology (eIT), Mankato, MN, USA, 2022. [CrossRef]
11. J. Zhang, M. Zhang, G. Qi, and S. Wu, "Reading recognition method of mechanical pointer meter based on machine vision," *Eng. Adv.*, vol. 3, no. 1, pp. 34–38, 2023. [CrossRef]
12. X. Li, C. Meng, X. Xiao, C. Yan, and Y. Lin, "Study on intelligent image recognition of non-linear short pointer SF6 meter readings," E3S Web of Conferences, Vol. 299, 2021, p. 03007.
13. Y. Sun, W. Li, Y. Zhao, L. Ren, and W. Jiang, "A method for image recognition of intersectant dual-pointer instrument," *Meas. Sens.*, vol. 18, p. 100224, 2021. [CrossRef]
14. L. Hou, Q. Zhang, and H. Qu, "Automatic reading recognition system of pointer meters using wsns with on-sensor image processing," *Eng. Res. Express*, vol. 3, no. 2, p. 025037, 2021. [CrossRef]
15. H. Nagarajan et al., "Automated analogue dial reading in cockpits," IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, Vol. 2022, 2022. [CrossRef]
16. W. Khan, D. Ansell, K. Kuru, and M. Amina, "Automated aircraft instrument reading using real time video analysis," IEEE 8th International Conference on Intelligent Systems (IS), Sofia, Bulgaria, 2016. [CrossRef]
17. F.-Y. Hsiao, F.-Y. Chang, P. Vida, B. C. Kuo, and P.-C. Chen, "Reading detection of needle-type instrument in a noisy environment using computer vision-based algorithms," *Multimedia Tool. Appl.*, vol. 82, no. 2, pp. 1749–1782, 2023. [CrossRef]
18. M. Zhang, Y. Zhou, and Q. Gu, "Design of image recognition algorithm for critical flight parameters in civil aircraft display system," *IOP Conference Series: Materials Science and Engineering*, vol. 751, no. 1, p. 012077, 2020.
19. J. M. Ricciardi, and C. Minwalla, "Non-intrusive flight test instrumentation using video recognition," AIAA Modeling and Simulation Technologies Conference, San Diego, California, USA, 2016. [CrossRef]
20. "NASA cockpit interruptions and distractions". [Online]. Available: https://asrs.arc.nasa.gov/publications/directline/dl10_distract.htm. [Accessed: February 18, 2024].
21. M. Adamski, M. Adamski, A. Adamski, and A. Szelmanowski, "Study of the Pilot's attention in the cabin during the flight," *J. KONES*, vol. 25, no. 3, pp. 17–24, 2018.
22. G. Ziv, "The need for eye tracking studies in helicopter pilots: A position stand," 1st International Workshop on Eye-Tracking in Aviation. Toulouse, Zurich, Switzerland: ETAVI, 2022.

23. S. Shappell, C. Detwiler, K. Holcomb, C. Hackworth, A. Boquet, and D. A. Wiegmann, "Human error and commercial aviation accidents: An analysis using the human factors analysis and classification system," *Hum. Factors*, vol. 49, no. 2, pp. 227–242, 2007. [CrossRef]
24. O. Lefrancois, N. Matton, G. Yves, V. Peysakhovich, and M. Causse, "The role of Pilots' monitoring strategies in flight performance," European Association for Aviation Psychology Conference, Cascais, Portugal, 2016.
25. S. Naeeri, and Z. Kang, "Exploring the relationship between pilot's performance and fatigue when interacting with cockpit interfaces," Institute of Industrial and Systems Engineers (IISE) Annual Conference, Norcross, Georgia, 2018.
26. F. Dehais, J. Behrend, V. Peysakhovich, M. Causse, and C. D. Wickens, "Pilot flying and pilot monitoring's aircraft state awareness during go-around execution in aviation: A behavioral and eye tracking study," *Int. J. Aerosp. Psychol.*, vol. 27, no. 1–2, pp. 15–28, 2017. [CrossRef]
27. D. Tappan, and M. Hempleman, "Image processing for data acquisition and machine learning of helicopter flight dynamics," International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV), 2016.
28. E. Tunca, H. Saribas, H. Kafali, and S. Kahvecioglu, "Determining the pointer positions of aircraft analog indicators using Deep Learning," *Aircr. Eng. Aerosp. Technol.*, vol. 94, no. 3, pp. 372–379, 2022. [CrossRef]
29. W. Khan, D. Ansell, K. Kuru, and M. Bilal, "Flight Guardian: Autonomous Flight Safety Improvement by Monitoring Aircraft Cockpit Instruments" *J. Aerosp. Inf. Syst.*, vol. 15, no. 4, pp. 203–214, 2018. [CrossRef]
30. J. M. Ernst, L. Ebrecht, and B. Korn, "Virtual cockpit instruments—How head-worn displays can enhance the obstacle awareness of helicopter pilots," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 36, no. 4, pp. 18–34, 2021. [CrossRef]
31. M. Alkhelaiwi, W. Boulila, J. Ahmad, A. Koubaa, and M. Driss, "An efficient approach based on privacy-preserving deep learning for satellite image classification," *Remote Sens.*, vol. 13, no. 11, p. 2221, 2021. [CrossRef]
32. GeoFS. [Online]. Available: https://www.geo-fs.com/. [Accessed: October 25, 2023].
33. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.*, vol. 9, no. 1, pp. 62–66, 1979. [CrossRef]
34. G. Jocher, A. Chaurasia, and J. Qiu, *YOLO by Ultralytics*, version 8.0.0 [Online]. Available: https://github.com/ultralytics/ultralytics. [Accessed: Oct. 25, 2023].
35. K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv Preprint ArXiv:1409.1556*, 2014.
36. S. H. Wang, M. A. Khan, and Y. D. Zhang, "VISPNN: VGG-inspired stochastic pooling neural network", *Computers, Materials and Continua*, vol. 70, no. 2, pp. 3081–3097, 2022. [CrossRef]
37. S. H. Wang, S. L. Fernandes, Z. Zhu, and Y. D. Zhang, "AVNC: Attention-based VGG-style network for COVID-19 diagnosis by CBAM," *IEEE Sens. J.*, vol. 22, no. 18, pp. 17431–17438, 2022. [CrossRef]
38. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018. [CrossRef]
39. F. Chollet, "Xception: Deep learning with depthwise separable convolutions," IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017. [CrossRef]
40. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, October 5–9, 2015.
41. C. Rother, V. Kolmogorov, and A. Blake, "'GrabCut': Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, 2004. [CrossRef]

Fatma Gümüş earned her BSc and MSc degrees in computer engineering from Istanbul University in 2013 and 2016, respectively. She completed her PhD degree at Yıldız Technical University in 2021. Currently serving as an assistant professor in the Computer Engineering Department at the Turkish Air Force Academy, National Defence University, her research focuses on speech processing, natural language processing, and vision-based deep learning.

Can Eyüpoğlu received the BSc degree with high honor in Computer Engineering and Minor degree in Electronics Engineering from İstanbul Kültür University, Türkiye in 2012, the MSc and PhD degrees with high honor in Computer Engineering from Istanbul University in 2014 and 2018, respectively. From 2019 to 2021, he was an Assistant Professor with the Computer Engineering Department, Turkish Air Force Academy, National Defence University, İstanbul, Türkiye. He is currently an Associate Professor and Head of Department in the Computer Engineering Department, Turkish Air Force Academy, National Defence University. He has published about 60 papers in various esteemed journals and conferences and has been serving as a member of the reviewer board in nearly 40 prestigious academic journals. He is also on the editorial board of some reputable journals. His current research interests include image processing, artificial neural networks, machine learning, bioinformatics, and data privacy.