Kafkas Universitesi Veteriner Fakultesi Dergisi
Journal Home-Page: http://vetdergikafkas.org
E-ISSN: 1309-2251

Kafkas Univ Vet Fak Derg
31 (3): 425-436, 2025
DOI: 10.9775/kvfd.2025.34130

Research Article

U-Net-Based Approaches for Biometric Identification and Recognition in Cattle

Pınar CİHAN [1] Ahmet SAYGILI [1 (*)] Muhammed AKYÜZLÜ [1] Nihat Eren ÖZMEN [1]
Celal Şahin ERMUTLU [2] Uğur AYDIN [2] Alican YILMAZ [2] Özgür AKSOY [2]

[1] Tekirdağ Namık Kemal University, Çorlu Engineering Faculty, Department of Computer Engineering, TR-59860 Tekirdağ - TÜRKİYE
[2] Kafkas University, Faculty of Veterinary Medicine, Department of Surgery, TR-36100 Kars - TÜRKİYE

[(*)] **Corresponding author:** Ahmet SAYGILI
Phone: +90 282 250 2376
E-mail: asaygili@nku.edu.tr

How to cite this article?

Cihan P, Saygılı A, Akyüzlü M, Özmen NE, Ermutlu CŞ, Aydın U, Yılmaz A, Aksoy Ö: U-Net-based approaches for biometric identification and recognition in cattle. *Kafkas Univ Vet Fak Derg,* 31 (3): 425-436, 2025. DOI: 10.9775/kvfd.2025.34130

Article ID: KVFD-2025-34130
Received: 20.03.2025
Accepted: 20.05.2025
Published Online: 26.05.2025

Abstract

Animal welfare is a factor that directly affects productivity and is one of the cornerstones of sustainable agriculture and animal husbandry practices. Traditional identification methods cause animal stress and create opportunities for theft and fraud. This is because conventional identification methods, unlike biometric methods, do not use the animal's natural features; therefore, the identity can be more easily copied or imitated. To minimize these problems and enhance animal welfare, this study proposes a computer-aided animal identification and recognition system using retina biometrics. In this study: i) Experts manually segmented 80 RGB cattle retinal images. ii) The images were augmented using various angles, generating 540 images. iii) An identification system was developed using the U-Net, SA-UNet, and U-Net++ deep learning models. iv) The performance of the developed identification system was measured for both the original and augmented datasets using the Dice coefficient and IoU. The study's findings show that the identification system's most successful model was U-Net (with a validation accuracy of 97.4%). The findings of this study demonstrate that cattle identification and recognition systems using retinal images were achieved with high accuracy rates. This study investigates retinal recognition and evaluates the performance of deep learning models for retinal identification, while also providing a publicly available expert-annotated ground truth dataset.

Keywords: Biometric identification, Biometric recognition, Cattle retina biometrics, Deep learning segmentation, U-Net architecture

Introduction

Biometric identification in animals, which provides unique identification/recognition of animals by utilizing their physical characteristics, has become an important tool in modern agriculture and livestock sectors [1]. These methods include fingerprint recognition, facial recognition [2], retina scanning [3–5], nose prints [6], body patterns [7], and DNA profiling [8]. These methods give each animal a unique identifier, facilitating easy detection in cases of loss or theft. Biometric identification in animals offers significant benefits not only in cases of loss but also in areas such as health monitoring, reproduction control, and feed consumption tracking [9].

One of the main advantages of biometric identification is that it does not negatively impact animal welfare. Traditional methods have been used for animal identification and recognition up until now. These traditional methods include ear notching, ear tattoos, branding, freeze marking, and ear tagging [10]. Traditional methods can cause stress in animals, increase the risk of infection, and even subject them to physically traumatic procedures.

Ear tagging is still widely used for animal identification and recognition [11]. This method involves attaching unique numbered tags to the animals' ears. However, ear tagging has numerous disadvantages. For instance, the process can cause infections or irritation in the animals' ears. Animals may experience stress, negatively affecting their production efficiency. If the animal moves its head during tagging or an issue arises, it can result in can result in torn ears. Additionally, animals can snag their tags on fences or other objects, leading to ear injuries. Moreover, the ear tagging method has environmental drawbacks. Tags frequently fall off, contributing to plastic waste and environmental pollution. The loss of tags necessitates re-

tagging, which incurs additional costs and subjects the animals to stress again. Tags can also be easily removed or copied, leading to theft and fraud [12]. Consequently, insurance premiums are high, prompting animal owners to avoid insurance. These economic losses associated with ear tagging negatively affect animal owners and the national economy. Therefore, it is crucial to transition to more effective and modern identification methods to eliminate these problems.

Unlike biometric systems, other tracking methods monitor devices rather than animals. Modern technologies like biometric identification provide a safer and more effective way to identify and recognize animals individually. For example, biometric methods such as retina scanning can verify animals' identities using their unique physical characteristics. As with ear tags, these methods can be performed without subjecting animals to external influences and stress. Additionally, since these methods involve scanning body parts, they do not produce waste and thus do not contribute to environmental pollution. Biometric methods cannot be copied, preventing fraud, and there is no risk of loss or deformation over the animal's lifetime. These methods significantly reduce stress levels and protect animal welfare [13–15]. Transitioning to modern identification methods enhances animal welfare, promotes healthier and more efficient livestock practices, reduces the economic burden on animal owners, and supports an environmentally friendly approach, ultimately benefiting the national economy [16]. To promote these methods, necessary infrastructure and training support should be provided, and awareness should be raised within the industry.

In this study, a system for animal identification and recognition from retina biometry was developed using digital image processing methods. When the literature examines studies based on retinal images, it shows a limited number of studies and that these studies generally use ready-made software [17-24]. The retinal vascular pattern is a biometric identifier present from birth, remaining unchanged throughout the animal's life and incapable of being imitated, making it the most secure biometric identification method [25]. This pattern can differ even between twins, clones, and the eyes of the same animal [26]. Obtaining retinal images is painless, easy, reliable, and cost-effective. Other biometric structures, such as fingerprints, faces, palms, and irises, can be altered through plastic surgery, making them less secure than retinas. Numerous researchers have supported the use of retina imaging technology for identifying farm animals [4,19,22,24,27].

To transition to biometric systems, it is essential to first record the biometric data of animals in databases. However, few countries have started identification and recognition studies using biometric markers. Moreover, collected biometric data is kept private. This study investigates retinal recognition and evaluates the performance of deep learning models for retinal identification. Previous studies have focused solely on animal identification without using deep learning methods. In addition to identification, we developed a recognition system in this study. This allows for comparing any given retinal image with all other images using a matching score, determining whether the image belongs to the same animal as the one with the highest match. Data was first collected in Türkiye using retina biometrics and deep learning to identify and recognize cattle. Labeled images are essential for training deep learning methods. Therefore, 80 retina images were annotated by experts to create a ground truth dataset. This number is relatively high, and the dataset manually labeled and frequently used for humans in the literature consists of 40 retinas (DRIVE Dataset). After training with this dataset, the identification performance of U-Net, SA-UNet, and U-Net++ deep learning models was compared. Finally, the recognition system performance of the U-Net model, which showed superior results in cattle identification, was evaluated.

The main contributions of this article are listed below:

- Addressing a gap in literature focused solely on identification by developing a recognition system.
- Demonstrating the effectiveness of deep learning models for retinal identification and recognition.
- To compare the performances of U-Net, SA-UNet, and U-Net++ deep learning models in vessel segmentation.
- To compare the performances of BRISK (Binary Robust Invariant Scalable Keypoints), FAST (Features from Accelerated Segment Test), HARRIS (Harris Corner Detection), SIFT (Scale-Invariant Feature Transform), and SURF (Speeded-Up Robust Features) feature extraction methods.
- To publicly share an expert-annotated ground truth dataset of 80 retinal images.

Material and Methods

Ethical Statement

The study was approved by the Kafkas University Animal Experiments Local Ethics Committee in Turkey (Protocol number: KAÜ-HADYEK/2025-018).

Database Used

The animal material of this study was obtained by taking retinal images from 300 cattle brought to the Faculty of Veterinary Medicine of Kafkas University, Türkiye. The datasets used in this study were collected within the scope of previously published studies [3-5]. The breed, age, sex

and ear tag numbers of the cattle brought to the clinic were recorded. During the data collection phase, the sick animals were first examined in general clinically and recorded if any disease was detected. Then, images were collected in a closed environment using the Optomed Smartscope digital fundus camera, with at least two images from both the right and left eyes. The Optomed Smartscope digital fundus camera was used only to collect retinal images. Deep learning and image processing studies were performed with the models created within the scope of this article. 80 of the 300 cattle collected in this study were manually segmented and used by experts. The reason for selecting 80 images for manual segmentation is that the process is labor-intensive and requires expertise. In the literature, the maximum number of manually segmented retinal images for humans is 40 (DRIVE Database), highlighting the challenge of this task.

Original RGB retina images captured with an Optomed portable fundus device were archived in JPG format at 1536×1152 resolution. 80 RGB fundus images, manually segmented by experts, were resized to 512×512 while preserving the aspect ratio. The manually segmented binary images were saved in PNG format. The segmented image dataset has been publicly shared on Kaggle *(https://www.kaggle.com/datasets/animalbiometry/cattle-retinal-fundus-groundtruth).*

Additionally, augmentation techniques were applied to the images to measure the models' generalization performance. During field capture of retina images from animals, the device's angle of grip could alter the angle of the retina image. Therefore, in this study, the dataset was augmented by rotating 80 original retina images by 0°, 30°, 60°, 90°, 120°, 150°, and 180°. After augmentation, a total of 560 augmented retina images were obtained.

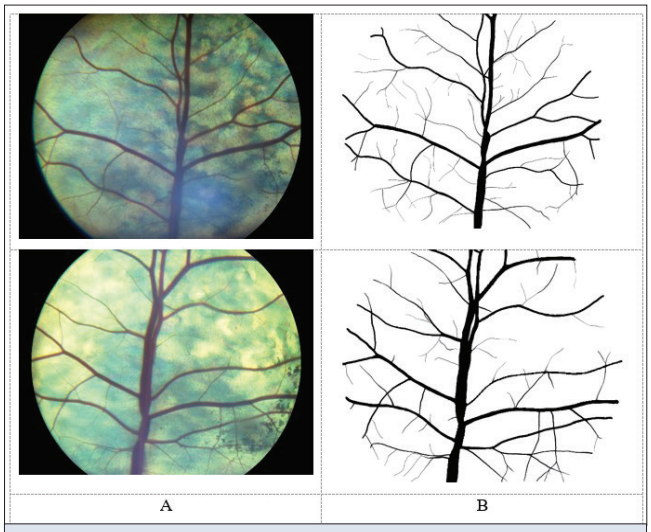

Fig 1. A) Original RGB image, **B)** manually labelled (expert-annotated) image

Fig. 1 presents an example of an RGB retina image and its labelled BW image. During the training phase of deep learning models, both original (n=80) and augmented (n=560) retinal images were utilized. Additionally, an independent set of 1,206 distinct RGB retinal images was employed to evaluate the models' performance.

Retinal Vessel Segmentation And Feature Extraction

In this study, deep learning methods U-Net, SA-UNet, and U-Net++ were used to segment the retinal vascular structure of cattle. After vessel segmentation, features in the images were extracted using BRISK, FAST, HARRIS, SIFT, and SURF methods.

Deep Learning Models

The U-Net, SA-UNet, and U-Net++ deep learning models were utilized for retinal vessel segmentation. Each offers distinct approaches to enhance segmentation performance.

U-Net Model is a convolutional neural network developed for biomedical image segmentation [28]. The fundamental idea behind U-Net is to spatially reduce the image feature map size in the network to store only important features and discard less valuable data, then create a bottleneck to learn significant features and restore them to the original size. The architecture consists of encoder and decoder blocks. The contractive branch (encoder) uses traditional convolution to down-sample the image's representation and produce a compressed feature representation of the input image. The expansive branch (decoder), complementary to the contractive branch, uses up-sampling methods like transpose convolution to ensure the processed output is the same size as the input. In *Fig. 2*, each blue box in the U-Net architecture corresponds to a feature map. Numbers written above the boxes represent channel numbers, and those in the bottom left corners indicate dimensions. White boxes denote copied feature maps, and arrows indicate different operations. The network structure of the U-Net algorithm used for retina vessel segmentation is schematically illustrated in *Fig. 2*.

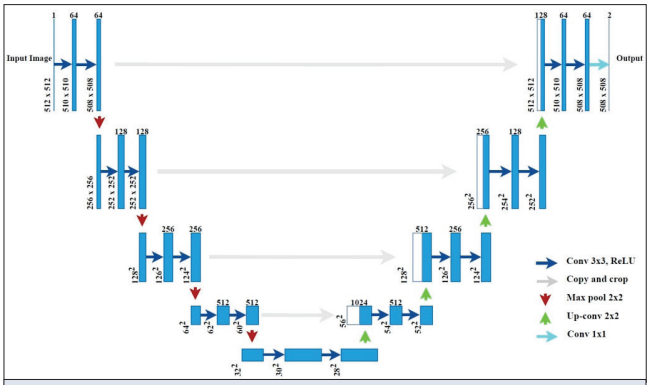

Fig 2. Diagram of U-Net architecture

SA-UNet Model is widely used in medical image segmentation, various variants have been proposed to achieve even better performance. These variants have improved performance but have made the network more complex and less interpretable. To address these issues, Spatial Attention U-Net (SA-UNet) was proposed [29]. In the SA-UNet architecture, convolutional blocks are replaced with structured convolution blocks integrating DropBlock and BN (batch normalization). This enhances the network's representation ability by focusing on vascular features and suppressing insignificant features by adding a few extra parameters. The SA-UNet network structure used for retina vessel segmentation is schematically illustrated in *Fig. 3*. The difference between vascular and non-vascular features in retina fundus images, especially in small and marginal vascular areas, is not distinct.

U-Net++ Model, the encoder captures high-level features from the input image through a series of convolutional and pooling layers. At the same time, the decoder uses

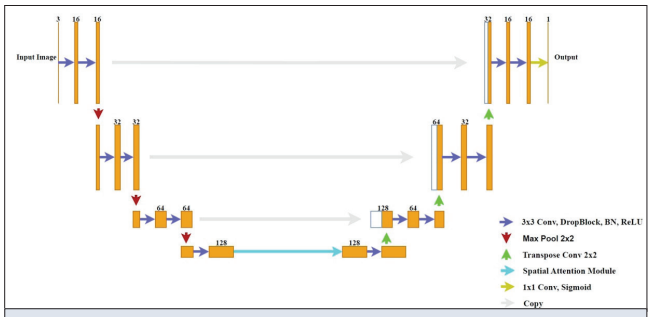

Fig 3. Diagram of SA-UNet architecture

up-sampled representations of these features to generate a dense segmentation map. However, there can be a semantic gap between encoder and decoder features, which may challenge the decoder to reconstruct fine details and produce accurate segmentation. UNet++ addresses this semantic gap by introducing the concept of nested skip pathways. Multiple skip connections are placed between the encoder and decoder blocks at different resolutions. These connections allow the decoder to access and fuse both low-level and high-level features from the encoder, enabling a more comprehensive understanding of the image in finer detail [30]. The UNet++ network structure used for retina vessel segmentation is schematically illustrated in *Fig. 4*.

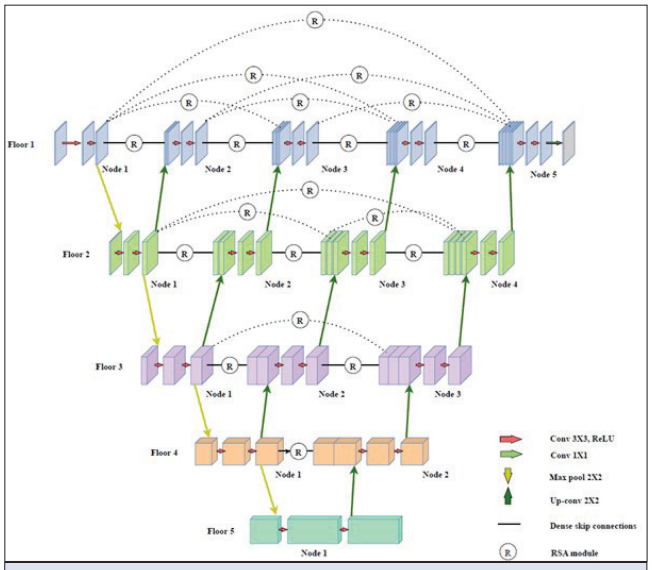

Fig 4. Diagram of UNet++ architecture

Feature Extraction Methods

BRISK, FAST, HARRIS, SIFT, and SURF methods were used for feature extraction in retinal images, and this section provides their descriptions, and the parameters used in the study.

Binary Robust Invariant Scalable Keypoints (BRISK) is a feature extraction algorithm in computer vision and image processing. The processing steps of the BRISK algorithm are as follows:

1. ***Corner Detection:*** BRISK detects specific corner points in the image using a scale space (pyramid structure).
2. ***Orientation Calculation:*** Each key point is assigned to an orientation value based on the bright variations of the surrounding pixels.
3. ***Binary Descriptor Creation:*** Binary descriptors are created using the bright differences of pixel pairs around the key point.
4. ***Matching:*** The binary descriptors are compared using Hamming distance to determine matches.
5. ***Non-Maximum Suppression:*** This method identifies the strongest points to select the most prominent corners.
6. ***Scalability:*** BRISK offers a scalable structure to accurately detect objects of different sizes.

Features from Accelerated Segment Test (FAST) aims to identify prominent distinctive points (corners) in images. Corners are points that contain important information about the geometric structure of an object and are used in many applications, such as object recognition and image matching. The steps of the FAST algorithm are as follows:

1. ***Corner Candidate Detection:*** The algorithm detects corner candidates by comparing a specific pixel with its surrounding pixels. The pixels are evaluated based on the bright differences along a circular contour.
2. ***Circle Definition:*** A circle with a radius of 16 pixels is defined around each corner candidate. The pixels

within this circle are classified as brighter, darker, or similar in brightness compared to the central pixel.

3. ***Threshold Value Check:*** The pixels on the circle must differ from the central pixel by a certain threshold value. If at least 12 pixels are found to be either brighter or darker than the center pixel, that pixel is considered a corner.
4. ***Accelerated Check:*** The algorithm does not evaluate all 16 pixels in the circle to determine the presence of a corner. Instead, it initially examines the diagonal pixels (e.g., pixels 1, 5, 9, and 13). The evaluation process is terminated if these four pixels fail to satisfy the corner criteria.
5. ***Non-Maximum Suppression:*** Among the detected corners, the strongest ones are selected for feature extraction using the non-maximum suppression method.

The HARRIS method is a gradient-based corner detection algorithm. The steps for extracting features are as follows:

1. ***Gradient Calculation:*** Compute the gradients in the x and y directions around each pixel using Sobel or other edge detection operators to find the rate of brightness change.
2. ***Harris Matrix Formation:*** Using the gradients, construct a structure called the Harris matrix, which helps determine if a pixel is a corner.
3. ***R Score Calculation:*** Calculate an "R" score for each pixel to classify it as a corner, edge, or flat region.
4. ***Corner Detection:*** If the R score exceeds a certain threshold, the pixel is marked as a corner; positive R indicates a corner, while negative R indicates an edge or flat area.
5. ***Non-Maximum Suppression:*** Filter the corners with high R scores using non-maximum suppression to identify and extract the strongest corners as features.

The introduced Scale Invariant Feature Transform (SIFT) method is an image descriptor for image-based matching proposed by David G. Lowe in 2004. The descriptor is used for numerous purposes in computer vision related to point-matching different views of a 3D scene and view-based object recognition. The SIFT descriptor is invariant with translations, rotations, and scaling transformations in the image space and is invariant with perspective transformations and illumination changes [31].

The SIFT algorithm consists of four steps [32]:

1. ***Recognition of key points with scale space:*** Points of interest (key points) are detected by blurring images at different scales with a Gauss filter and creating different images (DoG).
2. ***Improvement of key points according to contrast and edge thresholds:*** The detected points of interest are positioned more precisely, and points with low contrast or located at the edges are eliminated.
3. ***Assigning the direction of each key point:*** Each key point is assigned a direction by calculating the gradient directions around it. This ensures that the features remain stable against rotations.
4. ***Creating feature definitions:*** Using the gradient magnitude and direction information around each key point, feature vectors are created. These vectors are used for comparison and matching.

SIFT is used to detect points of interest from a grey-level image, giving a summative description of local image structures in a local neighborhood around each point of interest. This descriptor is used to match corresponding points of interest between different images. SIFT is used for studies such as object categorization, texture classification, image alignment, and biometrics [33].

The Speeded Up Robust Features (SURF) method is an algorithm used for detection and recognition in the fields of image processing and computer vision [34]. SURF is a scale-invariant feature detector based on the Hessian matrix. Instead of using a different metric to select position and scale, a Hessian detector is used for both. The Hessian matrix is a square matrix of second-order partial derivatives of a numerical field and is roughly analogous to using a series of box-type filters. The main interest of the SURF approach is that it quickly computes operators using box filters, thus enabling real-time applications such as tracking and object recognition [35].

The SURF algorithm consists of four main parts:

- Integral image generation
- Hessian detector
- Descriptive orientation assignment
- Creating an identifier

The integral image $I_{\sum}$ is calculated with Equation 1.

$$I_{\sum}(x,y) = \sum_{i=0}^{i\leq x}\sum_{j=0}^{j\leq y} I(x,y) \quad (1)$$

Here I is the input image. The time required for the calculation is invariant to change in size. The Hessian matrix is given in Equation 2.

$$H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix} \quad (2)$$

L_{xx}, is the convolution of the second-order Gaussian derivative with the image at x = (x, y). L_{yy} *ve* L_{xy} are treated similarly. The determinant of this matrix is used to find points of interest. Each point is compared to eight

points on the local scale, nine points on the upper scale, and nine points on the lower scale.

Table 1 provides the parameters used in the study to obtain features from retinal vessels using the BRISK, FAST, HARRIS, SIFT, and SURF methods.

Performance Measures

This study used accuracy, validation, and loss metrics to measure the training performances of U-Net, SA-UNet, and U-Net++ models. The 5-CV technique was used when training the models [4] FAR, FRR, precision, recall, and accuracy metrics were used to measure the recognition performance of the learning models, and the matching score approach was used to measure the identification performance of the models.

Table 1. *Model parameters of feature extraction methods*

Method	Parameter/Value
BRISK	thresh = 0 octaves = 3 patternScale = 1.0f
FAST	nonmaxSuppression = False threshold = 10
HARRIS	maxCorners = 1000 qualityLevel = 0.01 minDistance = 10 blockSize = 5 k = 4 mask = None
SIFT	nfeatures = 3000 nOctaveLayers = 30 contrastThreshold = 0.04 edgeThreshold = 1000 sigma = 1.6
SURF	hessianThreshold = 100 nOctaves = 4 OctaveLayers = 30 Extended = False upright = False

Accuracy is the ratio of correctly matched retinal images to the total number of retinal images, calculated as in Equation 3. This metric shows how accurately the model makes predictions during training.

$$\text{Accuracy} = \frac{\#correctly\ matched\ retina\ images}{\#total\ retina\ images} \quad (3)$$

Validation is the evaluation made on images that have not been used in the training data set before. It shows the model's success on new data (that it has not seen before). This doesn't happen based on the data the model sees during training, so predicting how well the model can perform on real-world data is important. Its formulation is the same as accuracy; only the data set changes.

Loss: the BCEDiceLoss function was used to calculate losses in cattle identification. BCEDiceLoss consists of the combination of Binary Cross Entropy (BCE) and Dice loss. Binary cross-entropy loss (BCE) is used to evaluate the probability that a pixel belongs to a particular class. On the other hand, Dice loss is mainly used in segmentation tasks and is known for determining object boundaries more effectively. Dice loss evaluates performance by measuring the agreement between pixels predicted by the model and actual labels. BCEDiceLoss combines the advantages of these two functions, allowing the model to increase pixel-wise accuracy and detect object boundaries more precisely. Thanks to this integration, the model's overall performance is improved, and the cattle retinal identification process is aimed at being more efficient and accurate. BCEDiceLoss function is calculated as in Equation 4.

$$\text{BCEDiceLoss}=\text{BCELoss}+(1-\text{DiceLoss}) \quad (4)$$

The purpose of the recognition system is to check whether the cattle whose retina has been scanned are in the database. According to the optimum threshold value determined here, whether the animal is in the database or not is questioned. The study defined the value at which the FAR metric was minimized as the optimum threshold value. According to the determined threshold value, the recognition performance of U-Net, SA-UNet, and U-Net++ models was measured with FAR, FRR, precision, recall, and accuracy metrics.

False Acceptance Rate (FAR) is a metric that expresses the rate at which cattle not in the database are mistakenly accepted as being in the database when checking whether the cattle whose retinas have been scanned are in the database. This ratio is used to evaluate the reliability and accuracy of the recognition system. The FAR formula is given in Equation 5.

$$FAR = \frac{FP}{FP + TN} \quad (5)$$

False Rejection Rate (FRR) is a metric that expresses the rate at which cattle in the database are mistakenly rejected as if they were not in the database while checking whether the cattle whose retina was scanned are in the database. This ratio is used to evaluate the accuracy and reliability of the recognition system. FRR is calculated as the rate at which the system cannot recognize correctly registered cattle, as in Equation 6.

$$FRR = \frac{FN}{FN + TP} \quad (6)$$

Precision refers to the ratio of cattle recognized by the system as existing in the database, among the cattle actually in the database, when checking whether the cattle whose retinas have been scanned are in the database. In other words, it is the ratio of the system's true positive

predictions to the total positive predictions and is calculated as in Equation 7.

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

Recall refers to how many of the cattle in the database are correctly recognized by the system when checking whether the cattle whose retinas have been scanned are in the database. It is the ratio of true positive predictions to total true positives and is calculated as in Equation 8.

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

The primary purpose of the identification system is to compare the retinal image of the tested cattle with all cattle retina images in the database to obtain a match score between retina pairs. The match score gives us the percentage of how many distinctive feature points of the two compared images overlap. The identification process is successful if the images with the highest match score belong to the same cattle. In this case, the identity of the cattle the system tests are determined by finding the closest match in the database. The match score is calculated using the formula in Equation 9.

$$Matching\ Score = \sqrt{\frac{(matchingPoints)^2}{(pointsImg1)*(pointsImg2)}} \times 100 \quad (9)$$

MatchingPoints represents the number of matching feature points between the reference and test images. pointImg1 indicates the total feature points in the reference image, and pointImg2 indicates the total feature points in the test image.

Results

This study used U-Net, SA-UNet, and UNet++ deep learning models for retinal biometric identification and recognition. Animal identification and animal recognition are distinct and equally important processes. Animal recognition focuses on determining whether an unknown animal exists in the database. In contrast, animal identification informs the user which animal matches the unknown animal in the database. To increase the efficiency of livestock farming, particularly in large-scale animal farms, it is crucial to implement identification and recognition processes. During the training phase of the models applied for identification and recognition, the parameters listed in *Table 2* were used. These parameters were determined based on experimental studies.

The 5-fold cross-validation method was used when training the models [36]. Each part was used for testing, while the remaining was reserved for training. The model's accuracy was generalized by taking the average of these five folds. Average Dice Similarity Coefficient (Dice) and Intersection Over Union (IoU) were used to quantitatively evaluate the performance of the methods and compare them with others. In *Table 3*, IoU and dice are calculated for two data sets (original and augmented) of each model and presented next to the ± deviation value.

Table 2. Training parameters of U-Net, SA-UNet, and UNet++ models

Parameter	Value
Optimizers	Adam
Loss	BCEDiceLoss
Size	512 x 512
K_Folds	5
Batch_size	2
Epochs	300
Learning Rate	0.001
LR Scheduler	ReduceLROnPlateau
Scaler	GradScaler

Table 3 compares different models' performance using Dice and IoU (Intersection over Union) metrics. Dice Score is a metric that measures the similarity of two clusters. Values closer to 1 indicate that the model shows better segmentation performance. IoU (Jaccard Similarity) measures segmentation success by evaluating the intersection and union ratio. High IoU values suggest that the model distinguishes the target object better. As can be seen from *Table 3*, when the data augmentation method is used, an improvement is observed in the Dice and IoU values of all models. In particular, the IoU and Dice scores of the U-Net and SA-UNet models showed the highest performance.

Table 3. Quantitative results of the models for the animal identification system (mean ± SD)

Dataset	Model	Dice	IoU
Original	U-Net	0.977±0.001	0.954±0.001
	SA-UNet	0.976±0.001	0.954±0.001
	UNet++	0.976±.001	0.953±0.002
Augmented	U-Net	**0.983±.000**	**0.966±0.001**
	SA-UNet	**0.983±0.000**	**0.966±0.001**
	UNet++	0.982±.001	0.965±0.001

As seen in *Table 3*, U-Net and SA-UNet models are the most successful models in retinal vessel segmentation, with a success rate of 0.983±0.000 from the Dice method and 0.966±0.001 from the IoU method. The U-Net++ model showed a very close but lower performance than the other models. When the performances of these models were compared in terms of accuracy, verification, and loss metrics, the results in *Table 4* were obtained.

Table 4. *Comparison of performances of U-Net, SA-UNet, and U-Net++ models in identification*

Dataset	Model	Training Acc (%)	Test Acc (%)	Loss (%)
Original	U-Net	96.9±0.2	96.6±0.2	6.9±0.3
	SA-UNet	96.9±0.1	96.6±0.1	6.9±0.3
	U-Net++	97.0±0.1	96.6±0.1	6.9±0.1
Augmented	U-Net	**97.8±0.1**	**97.4±0.0**	**5.1±0.1**
	SA-UNet	97.8±0.1	97.4±0.1	5.1±0.1
	U-Net++	97.7±0.1	97.3±0.1	5.2±0.1

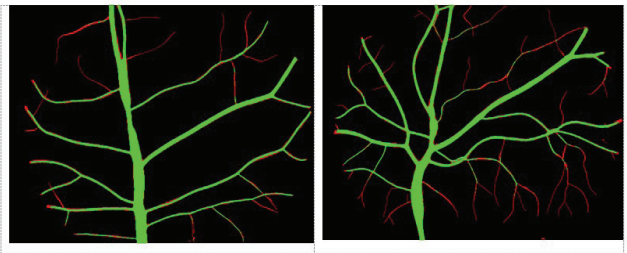

Fig 5. Ground truth U-Net matching result **A)** sample 1, **B)** sample 2

Table 4 compares the performance of U-Net, SA-UNet and U-Net++ models during training. Training accuracy represents the model's accuracy on the training data, while validation accuracy indicates its performance on the validation data. The loss function reflects the model's error rate, with lower values indicating better optimization. When data augmentation is applied, an increase in both training and validation accuracy is observed across all models, while the loss function values decrease. This demonstrates that data augmentation enhances the model's learning capacity and generalization ability.

In conclusion, when considering *Table 3* and *Table 4* together, the U-Net model is the most successful in vessel segmentation, achieving a validation accuracy of 0.974±0.000.

Fig. 5 presents the intersection of the vessel patterns segmented by the U-Net model with the manually segmented vessel patterns for the most successful model. In these retina images belonging to two different animals, the red indicates manually annotated vessels, while the green shows the sections matched by the U-Net model.

The developed computer-aided cattle identification and recognition system aims to prevent theft and fraud while ensuring animal welfare. The feasibility of such systems continues to be explored, as research in this area is ongoing. The findings obtained in this study demonstrate that the identification and recognition system can be successfully implemented. In this study, after evaluating the training performances of the models for identification in *Table 3* and *Table 4*, a recognition system based on U-Net was developed. The augmented dataset was used in the recognition system because higher success was achieved with this dataset. The advantage of this dataset is its higher generalization capability due to the inclusion of images at different angles. The performance of five different feature extraction methods (BRISK, FAST, HARRIS, SIFT, and SURF) used in the developed U-Net-based recognition system was measured, and the results are presented in

Table 5. *Performance of the animal recognition system for the augmented dataset*

Method	FAR (%)	FRR (%)	Precision (%)	Recall (%)	Accuracy (%)	Threshold	Exe.Time (s)
BRISK	18.52	17.49	97.84	82.51	82.42	30.50	1.80
FAST	10.34	12.84	98.75	87.16	87.40	22.60	4.80
HARRIS	11.03	11.40	98.33	88.60	88.64	36.90	0.24
SIFT	20.51	20.91	99.14	79.09	79.10	33.40	6.00
SURF	30.77	30.76	99.03	69.24	69.24	25.50	216.00

Table 6. *Confusion matrix of the recognition system (for 1206 test retina images)*

BRISK			
Actual Label	(-)	88	20
	(+)	192	906
		(-)	(+)
		Predicted Label	

FAST			
Actual Label	(-)	104	12
	(+)	140	950
		(-)	(+)
		Predicted Label	

HARRIS			
Actual Label	(-)	121	16
	(+)	129	940
		(-)	(+)
		Predicted Label	

SIFT			
Actual Label	(-)	31	8
	(+)	244	923
		(-)	(+)
		Predicted Label	

SURF			
Actual Label	(-)	18	8
	(+)	363	817
		(-)	(+)
		Predicted Label	

Table 5. In addition, the execution times of the methods are also shown in *Table 5*. As can be seen from *Table 5*, the fastest model is HARRIS with 0.24 sec.

As shown in *Table 5*, feature extraction from the retinal vessels was performed using the BRISK, FAST, HARRIS, SIFT, and SURF models. The threshold values in *Table 5* were determined experimentally and represent the optimal values that maximize the system's performance. These values indicate the degree of match between the retinal features of the scanned animal and those in the database. Specifically, suppose a value smaller than the threshold shown in *Table 5* is obtained when matching an animal's retina with the database. In that case, it indicates that the animal is not present in the database. For instance, in the model where features were extracted using the SIFT method segmented by the U-Net technique, a threshold level above 33.4 suggests that the scanned animal is present in the database, meaning it has been previously identified.

BRISK exhibits relatively high false acceptance (FAR: 0.1852) and false rejection rates (FRR: 0.1749). However, its precision is high at 0.9784, suggesting strong performance in minimizing false positives. Despite this, with a recall of 0.8251 and accuracy of 0.8242, BRISK demonstrates average overall performance compared to other methods.

FAST presents one of the lowest false acceptance (FAR: 0.1034) and false rejection (FRR: 0.1284) rates. Its precision (0.9875) and recall (0.8716) are high, and with an accuracy of 0.8740, FAST outperforms BRISK, indicating superior recognition accuracy.

HARRIS also shows low FAR (0.1103) and FRR (0.1140), along with high precision (0.9833) and recall (0.8860). These metrics demonstrate that the method effectively minimizes false positive and false negative results, achieving an overall accuracy of 0.8864.

SIFT, while having slightly higher false acceptance (FAR: 0.2051) and false rejection (FRR: 0.2091) rates than other methods, achieves high precision at 0.9914, implying that false positives are rare. Nevertheless, its recall (0.7909) and accuracy (0.7910) remain average.

SURF displays the highest FAR (0.3077) and FRR (0.3076) values, indicating significant false acceptance and rejection levels. Although SURF's precision is excellent at 0.9903, its recall (0.6924) and accuracy (0.6924) are lower than the other methods.

Table 6 presents the confusion matrices obtained. The recognition process was performed using the matching scores given in Equation 9. *Table 6* shows correct and incorrect predictions for features obtained using the BRISK, FAST, HARRIS, SIFT, and SURF methods with these scores. In the confusion matrix, the (+) label shows the correctly identified ones, while the (-) label shows the incorrectly identified ones. *Table 6* shows that the number of animals correctly identified in the database is relatively high. However, the rate of incorrectly identifying a foreign animal not in the database is also higher. Overcoming this issue is believed to be achievable by increasing the number of animals in the database as much as possible. Additionally, having a more significant number and variety of retina images for the animals in the database will help address this issue.

In conclusion, the HARRIS method is the most successful model due to its low false acceptance and rejection rates and high precision and recall values. While FAST also demonstrates strong performance, HARRIS achieves

the highest accuracy and recall, making it the most effective method overall. SIFT, although showing excellent precision, has relatively lower recall and accuracy. BRISK and SURF exhibit lower overall success rates. Therefore, HARRIS emerges as the best-performing method in this study. These results revealed that the U-Net+HARRIS based recognition system can be used effectively.

Discussion

This study used deep learning methods for animal identification and recognition from retinal images. Our findings demonstrated high accuracy rates, suggesting that the developed system is suitable for practical applications in animal biometric identification.

Previous studies have also emphasized the potential of retinal images as a unique biometric identifier in livestock. For example, Barry et al.[17] and Rojas-Olivares et al.[24] reported that retinal imaging technology is effective in distinguishing individual sheep and lambs, respectively. Similarly, Gionfriddo et al.[18] demonstrated the feasibility of using retinal images for identifying individual dogs. These studies primarily relied on traditional image analysis or proprietary software, whereas our study utilized deep learning models, achieving significantly higher segmentation and identification performance.

In terms of segmentation, our U-Net model achieved a Dice score of 0.983±0.000 and an IoU score of 0.966±0.001 on the augmented dataset, indicating excellent vessel segmentation performance. Comparable research by Mustafi et al. [20] using retinal images for goat identification reported moderate segmentation accuracy without deep learning enhancements, highlighting the performance advantage of our approach. Regarding the recognition system, while previous studies such as Allen et al.[22] mainly evaluated retinal images through manual or semi-automatic feature matching, our deep learning-based system combined automatic vessel segmentation with feature extraction techniques (BRISK, FAST, HARRIS, SIFT, SURF), resulting in a recognition accuracy of 88.64% with the HARRIS method. This is a significant improvement over the conventional methods reported in earlier works [17,19,22].

One limitation of our study is the manual segmentation of retina images, which, although reviewed for accuracy, can introduce subjective variability. Similar concerns were noted in earlier works where manual annotation was used [17,24]. Future studies can benefit from larger datasets and the application of semi-automatic or fully automatic annotation methods. Another noteworthy contribution is the public release of our annotated dataset and source code, which contrasts with prior studies where data and methods were often proprietary [3-5,22]. This openness is expected to accelerate multidisciplinary research in animal biometrics.

The differences between this study and the existing literature are as follows:

- All studies have focused on identification, meaning they check whether the animal is in the database. In this study, we developed both identification and recognition systems.
- Almost all studies in the literature have performed recognition using the software embedded in retinal imaging devices. In this study, we implemented the steps of digital image processing individually.
- The retinal images collected in previous studies are private, while the dataset in this study is publicly available on Kaggle.
- In our previous work, we only developed an identification system using image preprocessing methods, while this study demonstrates that deep learning models are more successful.

The proposed biometric identification and recognition system has a high potential for industrial and commercial use. In the livestock sector, especially in large-scale farms and meat production facilities, animal identification and tracking are critical requirements. While traditional methods rely on physical identification tools such as ear tags or microchips, biometric systems offer a more reliable and tamper-proof alternative.

The proposed system can make important data such as vaccination history, health status and genetic information of animals accessible quickly and accurately. This can increase efficiency by improving disease management and lineage tracking processes. At the same time, it can provide great convenience in animal trade and pre-slaughter health checks and traceability requirements. Considering the increasingly stringent traceability and animal welfare standards of the European Union and other international markets, it is also possible for this system to create a competitive advantage in global trade. In addition, biometric identification has great potential in improving quality control processes in meat and dairy production, increasing food safety, and providing consumers with more transparent information. In today's world where interest in sustainable agricultural practices is increasing, adopting such technologies can contribute to the digitalization of the livestock sector and accelerate its integration into modern agricultural systems. Therefore, the proposed biometric identification and recognition system has an important practical application in terms of efficiency, security, and traceability in the livestock sector, beyond being just an academic research topic.

In conclusion, a deep learning-based computer-aided

animal identification and recognition system was developed using retinal images. A computer-aided identification and recognition system can significantly contribute to the agriculture and livestock sector, enhancing animal welfare. According to the study's findings, the U-Net deep learning-based system achieved an identification accuracy of 97.4%. The system's identification accuracy pertains to determining whether an animal whose retina has been scanned has been previously identified. The recognition accuracy of the developed system was 88.64%, with a precision of 98.33%. The lower accuracy value of biometric recognition than identification can be explained as follows.

Biometric identification involves determining the specific identity of an animal among all animals in the database, enabling the tracking of its vaccination, health, and lineage records. In contrast, biometric recognition simply determines whether an animal exists in the database without specifying its identity. Biometric recognition is the process of determining whether an animal is in the database. Large herds, especially in animals such as cattle that are raised in different places and can be sold before being slaughtered, can be difficult to trace the origins. For this reason, biometric recognition can easily determine which animals belong to you. However, false matches or misses may occur since the system only answers "yes" or "no" in this process. When working with large data sets, errors can occur due to false negatives (not found in the database) or false positives (assumed to be present even though not in the database). The error rate can increase, especially when animals with similar retinal patterns are involved. Tracking animal movements increases efficiency and contributes to preserving meat quality by reducing the risk of disease spread.

Publicly available ground truth datasets of manually segmented animal retina images are scarce, posing challenges for multidisciplinary research in animal husbandry. This study contributes to the increase in multidisciplinary research, enabling the development of more successful identification and recognition systems.

Declarations

Availability of Data and Materials: The dataset used in the study is publicly available at https://www.kaggle.com/datasets/animalbiometry/cattle-retinal-fundus-groundtruth. The source codes are available at https://github.com/muhammedakyuzlu/retinal-vessel-detection-identification-unet-variants

Funding Support: This work was supported by the Turkish Scientific and Technical Research Council-TÜBİTAK (Project Number: 121E349).

Ethical Approval: The study has been approved by the Institutional Animal Care and Use Committee of Kafkas University (KAÜ-HADYEK/2025-018).

Competing of Interest: The authors declared that there is no conflict of interest.

Declaration of Generative Artificial Intelligence: The author have declared that the article, tables and figure were not written/ created by AI and AI-assisted technologies.

Author Contributions: Writing - review & editing, Writing - original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Project administration, Funding acquisition, Conceptualization: PC; Writing - review & editing, Writing original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Conceptualization: AS; Visualization, Validation, Software, Methodology: MA, NEÖ; Data curation: CŞE, UA, AY, ÖA. All authors reviewed the results and approved the final version of the article.

References

1. Cihan P, Saygılı A, Özmen NE, Akyüzlü M: Identification and recognition of animals from biometric markers using computer vision approaches: A review. *Kafkas Univ Vet Fak Derg,* 29 (6): 581-593, 2023. DOI: 10.9775/kvfd.2023.30265

2. Mahato S, Neethirajan S: Integrating Artificial Intelligence in dairy farm management – biometric facial recognition for cows. *Inf Process Agric,* 2024 (Article in press). DOI: 10.1016/j.inpa.2024.10.001

3. Cihan P, Saygılı A, Akyüzlü M, Özmen NE, Ermutlu CŞ, Aydın U, Yılmaz A, Aksoy Ö: Extraction of cattle retinal vascular patterns with different segmentation methods. *Sakarya Univ J Comput Inf Sci,* 7 (3): 378-388, 2024. DOI: 10.35377/saucis...1509150

4. Saygılı A, Cihan P, Ermutlu CŞ, Aydın U, Aksoy Ö: CattNIS: Novel identification system of cattle with retinal images based on feature matching method. *Comput Electron Agric,* 221:108963, 2024. DOI: 10.1016/j.compag.2024.108963

5. Cihan P, Saygılı A, Şahin Ermutlu C, Aydın U, Aksoy Ö: AI-aided cardiovascular disease diagnosis in cattle from retinal images: Machine learning vs. deep learning models. *Comput Electron Agric,* 226:109391, 2024. DOI: 10.1016/j.compag.2024.109391

6. Chan YK, Lin CH, Ben YR, Wang CL, Yang SC, Tsai MH, Yu SS: Dog nose-print recognition based on the shape and spatial features of scales. *Expert Syst Appl,* 240:122308, 2024. DOI: 10.1016/j.eswa.2023.122308

7. Sharma A, Randewich L, Andrew W, Hannuna S, Campbell N, Mullan S, Dowsey AW, Smith M, Hansen M, Burghardt T: Universal bovine identification via depth data and deep metric learning. *Comput Electron Agric,* 229:109657, 2025. DOI: 10.1016/j.compag.2024.109657

8. Antil S, Abraham JS, Sripoorna S, Maurya S, Dagar J, Makhija S, Bhagat P, Gupta R, Sood U, Lal R, Toteja R: DNA barcoding, an effective tool for species identification: A review. *Mol Biol Rep,* 50, 761-775, 2023. DOI: 10.1007/s11033-022-08015-7

9. Neethirajan S: Recent advances in wearable sensors for animal health management. *Sens Biosensing Res,* 12, 15-29, 2017. DOI: 10.1016/j.sbsr.2016.11.004

10. Conill C, Caja G, Nehring R, Ribó O: The use of passive injectable transponders in fattening lambs from birth to slaughter: Effects of injection position, age, and breed. *J Anim Sci,* 80 (4): 875-879, 2002. DOI: 10.2527/2002.804919x

11. Ahmad M, Ghazal TM, Aziz N: A survey on animal identification techniques past and present. *Int J Innov Comput,* 1 (2): 27-32, 2022.

12. Awad AI: From classical methods to animal biometrics: A review on cattle identification and tracking. *Comput Electron Agric,* 123, 423-435, 2016. DOI: 10.1016/j.compag.2016.03.014

13. Ratha NK, Connell JH, Bolle RM: Enhancing security and privacy in biometrics-based authentication systems. *IBM Syst J,* 40 (3): 614-634, 2001. DOI: 10.1147/sj.403.0614

14. Lee Y, Filliben JJ, Micheals RJ, Phillips PJ: Sensitivity analysis for biometric systems: A methodology based on orthogonal experiment designs. *Comput Vis Image Underst,* 117 (5): 532-550, 2013. DOI: 10.1016/j.

cviu.2013.01.003

15. Goudelis G, Tefas A, Pitas I: Emerging biometric modalities: A survey. *J Multimodal User In,* 2, 217-235, 2008. DOI: 10.1007/s12193-009-0020-x

16. Clark B, Stewart GB, Panzone LA, Kyriazakis I, Frewer LJ: A systematic review of public attitudes, perceptions and behaviours towards production diseases associated with farm animal welfare. *J Agric Environ Ethics,* 29, 455-478, 2016. DOI: 10.1007/s10806-016-9615-x

17. Barry B, Corkery G, Gonzales-Barron U, Mc Donnell K, Butler F, Ward S: A longitudinal study of the effect of time on the matching performance of a retinal recognition system for lambs. *Comput Electron Agric,* 64 (2): 202-211, 2008. DOI: 10.1016/j.compag.2008.05.011

18. Gionfriddo JR, Lee AC, Precht TA, Powell CC, Marren KK, Radecki SV: Evaluation of retinal images for identifying individual dogs. *Am J Vet Res,* 67 (12): 2042-2045, 2006. DOI: 10.2460/ajvr.67.12.2042

19. Barron UG, Corkery G, Barry B, Butler F, McDonnell K, Ward S: Assessment of retinal recognition technology as a biometric method for sheep identification. *Comput Electron Agric,* 60 (2): 156-166, 2008. DOI: 10.1016/j.compag.2007.07.010

20. Mustafi S, Ghosh P, Mandal SN: RetIS: Unique identification system of goats through retinal analysis. *Comput Electron Agric,* 185:106127, 2021. DOI: 10.1016/j.compag.2021.106127

21. Barry B, Barron UG, Butler F, Ward S, McDonnell K: Verification of sheep identity by means of a retinal recognition system. *Trans ASABE,* 54 (3): 1161-1167, 2011. DOI: 10.13031/2013.37081

22. Allen A, Golden B, Taylor M, Patterson D, Henriksen D, Skuce R: Evaluation of retinal imaging technology for the biometric identification of bovine animals in Northern Ireland. *Livest Sci,* 116 (1-3): 42-52, 2008. DOI: 10.1016/j.livsci.2007.08.018

23. Alturk G, Karakus F: Assessment of retinal recognition technology as a biometric ıdentification method in Norduz sheep. **In,** *Proceedings of 11th International Animal Science Conference.* 20-22 October, Cappadocia, Turkey, 2019.

24. Rojas-Olivares MA, Caja G, Carné S, Salama AAK, Adell N, Puig P: Retinal image recognition for verifying the identity of fattening and replacement lambs. *J Anim Sci,* 89 (2): 2603-2613, 2011. DOI: 10.2527/jas.2010-3197

25. Marchant J: Secure animal identification and source verification, *JM Communications, UK*, 1:28, 2002.

26. Caja G, Ghirardi JJ, Hernández-Jover M, Garín D: Diversity of animal identification techniques: From 'fire age'to 'electronic age.' **In,** *Proceedings of 17th International Conference on Antiviral Research*. 02-07 May, Tuscon, Arizona, USA, 2004.

27. Rusk CP, Blomeke CR, Balschweid MA, Elliott SJ: An evaluation of retinal imaging technology for 4-H beef and sheep identification. *J Ext,* 44 (5): 449, 2006.

28. Ronneberger O, Fischer P, Brox T: U-net: Convolutional networks for biomedical image segmentation. **In,** *Proceedings of Medical image computing and computer-assisted intervention-MICCAI 2015: 18th international conference*, October 05-09, Munich, Germany, 234-241, 2015.

29. Guo C, Szemenyei M, Yi Y, Wang W, Chen B, Fan C: SA-Unet: Spatial attention U-Net for retinal vessel segmentation. **In,** *Proceedings of 2020 25th international conference on pattern recognition (ICPR)*, 10-15 January, Milan, Italy, 1236–1242, 2021.

30. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J: UNet++: A nested U-net architecture for medical image segmentation. **In,** *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop*, September 20, Granada, Spain, 2018.

31. Lowe DG: Distinctive image features from scale-invariant keypoints. *Int J Comput Vis,* 60, 91-110, 2004. DOI: 10.1023/B:VISI.0000029664.99615.94

32. Sujin JS, Sophia S: High-performance image forgery detection via adaptive SIFT feature extraction for low-contrast or small or smooth copy–move region images. *Soft Comput,* 28437-45, 2024. DOI: 10.1007/s00500-023-08209-6

33. Lindeberg T: Scale invariant feature transform, *Scholarpedia,* 7 (5):10491, 2012. DOI: 10.4249/scholarpedia.10491

34. Bay H, Tuytelaars T, Van Gool L: Surf: Speeded up robust features. **In,** *Proceedings of Computer Vision-ECCV 2006: 9th European Conference on Computer Vision*, May 07-13, Graz, Austria, 2006.

35. Bay H, Ess A, Tuytelaars T, Van Gool L: Speeded-up robust features (SURF). *Comput Vis Image Underst,* 110 (3): 346-359, 2008. DOI: 10.1016/j.cviu.2007.09.014

36. Cihan P, Özcan HK, Öngen A: Prediction of tropospheric ozone concentration with Bagging-MLP method. *Gazi J Eng Sci,* 9 (3): 557-573, 2023. DOI: 10.30855/gmbd.0705087