GENDER IDENTIFICATION FROM LEFT HAND-WRIST X-RAY IMAGES WITH A HYBRID DEEP LEARNING METHOD

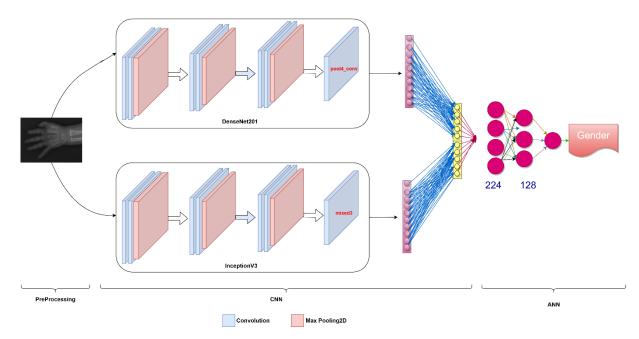[1, *]**Cüneyt ÖZDEMİR** , [2]**Mehmet Ali GEDİK** , [3]**Hüdaverdi KÜÇÜKER** , [4]**Yılmaz KAYA**

[1] *Siirt University, Engineering Faculty, Computer Engineering Department, Siirt, TÜRKİYE*
[2, 3]*Kütahya Health Sciences University, Evliya Çelebi Training and Research Hospital, Radiology Department, Kütahya, TÜRKİYE*
[4] *Batman University, Engineering Faculty, Computer Engineering Department, Batman, TÜRKİYE*
[1]cozdemir@siirt.edu.tr, [2]mehmetaligedik@gmail.com, [3]hudaverdi.kucuker@ksbu.edu.tr,
[4]yilmaz.kaya@batman.edu.tr

Highlights

- The success of transfer learning models on wrist images was compared.
- A Hybrid model was created from transfer learning models. While designing hybrid models, which layers should be preferred from transfer learning models was determined.
- Model's focus points while making gender discrimination were determined by observing heat maps, regions where gender discrimination is made on the wrist were identified.
- The effects of the model on age groups and the reasons for these effects were investigated.
- A high success rate of 96.3% was observed with proposed method for gender identification.

Graphical Abstract

Proposed gendering network architecture

*Corresponding Author: Cüneyt ÖZDEMİR, cozdemir@siirt.edu.tr

GENDER IDENTIFICATION FROM LEFT HAND-WRIST X-RAY IMAGES WITH A HYBRID DEEP LEARNING METHOD

[1, *] **Cüneyt ÖZDEMİR** ⓘ**,** [2]**Mehmet Ali GEDİK** ⓘ**,** [3]**Hüdaverdi KÜÇÜKER** ⓘ**,** [4]**Yılmaz KAYA** ⓘ

[1] *Siirt University, Engineering Faculty, Computer Engineering Department, Siirt, TÜRKİYE*

[2, 3]*Kütahya Health Sciences University, Evliya Çelebi Training and Research Hospital, Radiology Department, Kütahya, TÜRKİYE*

[4] *Batman University, Engineering Faculty, Computer Engineering Department, Batman, TÜRKİYE*

[1] cozdemir@siirt.edu.tr, [2] mehmetaligedik@gmail.com,
[3] hudaverdi.kucuker@ksbu.edu.tr, [4] yilmaz.kaya@batman.edu.tr

ABSTRACT: In forensic investigations, characteristics such as gender, age, ethnic origin, and height are important in determining biological identity. In this study, we developed a deep learning-based decision support system for gender recognition from wrist radiographs using 13,935 images collected from individuals aged between 2 and 79 years. Differences in all regions of the images, such as carpal bones, radius, ulna bones, epiphysis, cortex, and medulla, were utilized. A hybrid model was proposed for gender determination from X-ray images, in which deep metrics were combined in appropriate layers of transfer learning methods. Although gender determination from X-ray images obtained from different countries has been reported in the literature, no such study has been conducted in Turkey. It was found that gender discrimination yielded different results for males and females. Gender identification was found to be more successful in females aged between 10 and 40 years than in males. However, for age ranges of 2-10 and 40-79 years, gender discrimination was found to be more successful in males. Finally, heat maps of the regions focused on by the proposed model were obtained from the images, and it was found that the areas of focus for gender discrimination were different between males and females.

1. INTRODUCTION

In legal terms, gender identification is a primary component in forensic medicine practices related to the human body. Although there are bones such as skull bones and pelvis bones that facilitate gender determination, it is a medico-legal need to be able to determine gender from other bones in cases of increasing terror, murders, explosions, fires, and murder-accidental cases where the human body is dismembered and the above-mentioned bones are not found. In addition, gender determination is a necessity in murder-dismemberment and burning incidents against LGBT individuals.

Determining gender can also help us to understand the social and cultural origins of skeletal and human remains at archaeological sites [1].

Gender identification is also important for biological profiling, as is age and height, as well as for the evaluation of biological profile [2]. Human remains discovered in many forensic events may be fragmented or incomplete. In such a case, gender identification is important to obtain maximum information from the existing components of the skeleton [3]. Forensic anthropologist expertise is required when the forensic specialist cannot visually identify an individual's biological characteristics.

Accurate gender identification is often a crucial step in a forensic case as it can potentially narrow down the number of possible victims. However, determining gender can be challenging when dealing with incomplete or fragmentary remains, and morphological methods may not always be effective [4].

Since one of the most important biological markers for human identity in forensic cases is gender identification, researchers have tried to establish various osteometric standards by using different bones of the human skeleton. The skull is considered the most traditional and reliable part of the skeleton for

*Corresponding Author: Cüneyt ÖZDEMİR, cozdemir@siirt.edu.tr

gender identification. In cases of skull fragmentation, the evaluation of bones is quite difficult. Therefore, with the increase in the use of long bones for reliable gender identification, there has been a growing need for the use of alternative gender identification methods. It is stated that in forensic cases, the bones in the skeletal structure that are longer tend to be well-maintained and can be measured easily, making them dependable materials for determining an individual's gender. [5-9].

Gender identification is an integral part of constructing the biological profile. The inability to predict gender will further increase the difficulty in establishing estimates of race, weight, height, and age in a forensic setting. Prediction of biological traits is highly dependent on accurate gender determination.

Wrist radiographs are widely used in Turkey for the preparation of age determination reports requested by the courts in forensic cases. The "Sky Atlas", which was adapted from the atlas of Greulich Pyle (G-P), published by Şemsi Gök and his friends in 1969 and published in 1985, is used most frequently in forensic medicine practices in age determination. In addition, Greulich Pyle (G-P) atlas prepared according to the standards of western societies and Tanner Whitehouse (TW) method are also used [10].

There has been no previous study on gender determination from radius and ulna in Turkish society. Many of the studies in the literature have been done with traditional metric methods using ulna and radius images through the developed formulas.

In the literature research, it was observed that many different parts of the human were used for gender determination, but there were very few studies made from wrist radiographs. It has been observed that there are very few studies in which gender determination from wrist radiographs is made with artificial intelligence. For gender determination, the importance and success of wrist radiographs were measured. In this study, a deep learning-based decision support system using wrist bones was developed for gender determination. Deep learning is a branch of machine learning that utilizes artificial neural networks and algorithms modeled after the human brain's functioning to gain knowledge and insights from data. Similar to people's experience, it does a better job with a little tweaking each time to improve the result.

In this study, a decision support system has been developed by using deep learning methods in gender identification to overcome the problems that occur due to manual techniques. With the developed decision support system, it will be easier to get better results in a short time for gender determination, and quantitative and accurate evaluations will be provided. Furthermore, the regions where deep learning models focus on gender discrimination were determined using heat maps. In addition, the success of deep learning models in different age groups according to gender was compared, and the reasons for different success results according to gender were emphasized. In this study, 13935 left hand-wrist images were used to test the proposed hybrid model.

Some of the main contributions of this study are:

• The success of transfer learning models on wrist images was compared.

• A Hybrid model was created from transfer learning models. While designing hybrid models, which layers should be preferred from transfer learning models was determined.

• Model's focus points while making gender discrimination were determined by observing heat maps, regions where gender discrimination is made on the wrist were identified.

• The effects of the model on age groups and the reasons for these effects were investigated.

• A high success rate of 96.3% was observed with proposed method for gender identification.

2. LITERATURE STUDIES

There is limited research on gender identification from hand-wrist radiographs in the literature review. It has been observed that there are very few studies in the literature in which gender estimation is made using deep learning method from hand-wrist x-rays. In the literature research, it was observed that many different parts of the human were used for gender identification.

Many researchers have tried to determine gender by using different human body parts. Some of those include fingerprint [11, 13], iris ([14, 16] , palm print [17], chest X-ray [18], face[19, 21], hand

geometry [22], dental [23], skull [24], hair volume, face shape [25], ear [26], speech [27, 28], Coronavirus [48] and walking [29, 30]. These parts have been studied and these studies have obtained competitive results.

It has been observed that the number of studies that make gender identification from hand-wrist radiographs is relatively low. Only the research done by Yune et. al. [31] determines gender from bone, especially with artificial intelligence applications. Yune et al. [31] conducted their research on 10,607 images, 5459 females and 5148 males. They trained to fine-tune the VGG16 model, one of the transfer deep learning models, and then 95.9% accuracy was achieved in the tests performed on 1531 images.

Barnes et al. [32] used carpal bones on hand radiographs to determine gender in forensic cases in Thailand. Fifteen (15) measurements were made on seven carpals from 100 skeletons (50 males, 50 females) who died. As a result, they achieved a success rate of 91.8% in their study with statistical methods.

Studies have been conducted to determine gender using different human body parts. One of them is a study conducted by Vila-Bianco et al. [33], which determines gender from dental panoramic images. According to this research, three fully automated approaches based on deep learning architectures were compared in a database of 3400 dental panoramic images. The results provided 90% to 96% accuracy in patients older than 20 years.

Bewes et al. [34]; developed a deep learning system that makes gender identification from skeletal remains. The model trained with 900 skull images showed an accuracy of 95% with the GoogleNet model in determining gender.

Korot et al.[35] used deep learning models for gender identification from retinal fundus images. A model was trained on 84,743 retinal fundus photographs from the UK Biobank dataset. 86.5% success was reported in their studies.

Gornale et al.[36] stated that gender estimation is a process that helps to increase the security and efficiency of biometric systems, in their study. Gender identification using biometric features is mainly used to reduce search area lists, indexes and generate statistical reports. They achieved 99.9% success in their study with AlexNet, one of the deep learning models, with a dataset containing 15052 images. In addition, more successful results were obtained by modifying the transfer deep learning models.

In this study, the success of wrist X-rays, which is a different part of the body, was measured for gender prediction. It is the first to be conducted using deep learning models of gender identification with wrist x-rays collected on the Turkish population.

The data set KACRD (Kütahya Adult-Child Radiology Dataset) created by us was used to gender identification from hand-wrist X-ray images (approved by Kütahya University of Health Sciences, Ethical Committee Decision No: 221/07 dated 15.04.2021). The images in the data set were obtained from 4 different hospitals, collected from individuals aged 1-85 years. Dataset contains 13935 left hand-wrist x-rays, 7744 images from females, and 6192 images from males. Images were selected from the hospital archive among the left hand-wrist radiographs taken in patients admitted to the emergency department between 2013 and 2021. Images were obtained from different devices such as CR and digital radiography (DR). Due to the shooting in different centers and devices, some of the images' brightness and contrast settings differed from each other. Images are then converted to JPEG format and normalized. These images were excluded from the data set due to the very few images belonging to those younger than 2 years old and older than 79 years old. Studies were carried out on images aged 2-79 years. The distribution of images by gender in the data set is given in Table 1. An example of male and female x-ray images in the dataset is given in Figure 1.

Table 1. Distribution of datasets by gender

Gender	#images
Male	6191 (44.43%)
Female	7744 (55.57%)
Total Image	13935

Images of 27-year-olds of both genders

Male   Female

Figure 1. sample images from KACRD dataset

4. Method

4.1. Gender Detection Block Diagram

The deep learning architecture for determining gender from hand-wrist radiographs is outlined in this section. The approach proposed involves using 4 blocks, as depicted in Figure 2. The actions performed within each block are summarized.
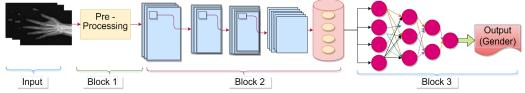

Figure 2. The general architecture of proposed deep learning method for gender identification

Block 1: It is the preprocessing stage. Necessary preprocesses are applied to all images in the dataset to remove brightness and contrast differences. Since the images were taken in different hospitals and from different angles, it was determined that many images' brightness and contrast settings were different from each other. The contrast limited adaptive histogram equalization method is used to solve this problem. After adjusting the contrast and brightness of the images in the preprocessing stage, the scaling process was performed. Deep learning methods involve complex mathematical operations, and reducing the size of the images can improve the speed of these operations [37]. The images were resized according to the target width and height, keeping the image aspect ratio the same without breaking them.

Block 2: After the preprocessing stage, the images are passed through the CNN block where various types of information and characteristics are extracted. At this block, InceptionV3 and DenseNet20 transfer deep learning methods were used.

Block 3: At this stage, classification is performed using the features from the previous block. A cause-effect relationship is determined between the features coming from the CNN block and the gender of the person.

4.2. Convolutional neural networks

In classical machine learning approaches, the extracted features were given to the classification algorithms after the features were extracted from the images with specific techniques. CNN is an artificial neural network that automates the feature extraction process. CNNs work with raw images as input, which are then processed through a multi-layer feedforward neural network composed of various types of layers such as convolution, activation, pooling, feature extraction with batch normalization, and a fully connected output layer. CNNs are designed to detect all types of information and features in images.

A key feature of deep learning is that overfitting can occur during training if the dataset is limited. Transfer learning methods are used to overcome this problem.

4.3. Transfer learning

CNN is better trained with large data sets. In certain fields, there may be a lack of data to train neural networks, leading to unsuccessful training and inadequate classification results. In such cases, using the transfer learning approach can be beneficial [38]. The performance of the neural network is affected by the success of the model and the number of operations in the network model. Therefore, using previously trained models and weights can improve the performance of the neural network, making the training process faster and more accurate. There are various transfer learning models available that can be used for different image classification or regression tasks. Transfer learning models are pre-trained models that can be used for different image classification or regression problems.

Morid et al.[39] conducted a large-scale study of medical images. A large-scale review of transfer learning in medical image analysis was conducted on 102 studies. According to this article, the most successful popular transfer learning models used for various anatomical regions are shown in Table 2.

Table 2. Popular transfer learning models used for different anatomical regions [39]

Anatomical Region	Model
Ultrasound, endoscopy, and skeletal system	Inception
For eye, skin, dental X-rays	VGGNet
For brain MRIs and breast X-rays	AlexNet
For chest X-rays	DenseNet

As can be seen from Table 2, Inception is shown as the successful transfer learning model on super skeletal systems. In this study, InceptionV3 and DenseNet20 models, which were frequently used in studies on medical problems, were used.

4.3.1. InceptionV3

The InceptionV3 consists of 42 layers. The main idea of the model is to learn multi-scale representations and replace small kernels with large kernels to reduce computational complexity and the total number of parameters [40]. This model was built using the GoogLeNet network architecture. The layout of the InceptionV3 network is illustrated in Figure 3.

The numbers "x3, x4, x2" in Figure 3 represent the number of times the process is repeated.
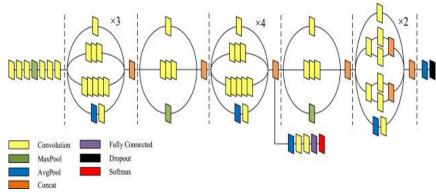
Figure 3. InceptionV3 network architecture [38]

4.3.2. DenseNet

DensetNet feed-forward connects each layer to every other layer, giving the properties of all previous layers as input to each layer. With this property, DenseNet enables the reuse of image features. The DenseNet network architecture is shown in Figure 4 [41]. It consists of 305 layers and 4 dense blocks. Each block has 128 and 32 convolutional blocks from 1x1 and 3x3 filters. Transition layers between two adjacent blocks with convolution and Max pooling (1 × 1 Convolution., 2 × 2 Max pooling) reduce the size of the feature maps. This is done to reduce bottleneck input feature maps and make convolution process more efficient.
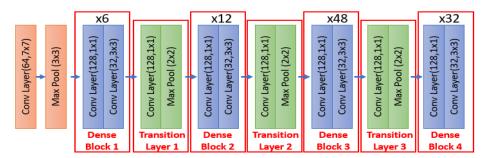

Figure 4. DenseNet201 layer architecture [41]

4.4. Proposed network architecture for gender identification

Table 4 shows the most successful results were obtained with the InceptionV3 and DenseNet201 models as a result of the experimental studies. For this reason, the layers that produce the best features in these 2 models were determined and a hybrid model consisting of these layers was designed. The designed hybrid model is shown in Figure 5. By looking at the activations in each layer of the InceptionV3 and DenseNet201 deep transfer models, the activations in the layers are visualized. These operations were performed for all layers. The visualized activations in each layer were then visually examined, and deep metrics were taken from the layer before the layer where the images started to deteriorate. Gender recognition was performed by combining deep metrics from InceptionV3 and DenseNet201 and giving them to the Fully Connected layer.
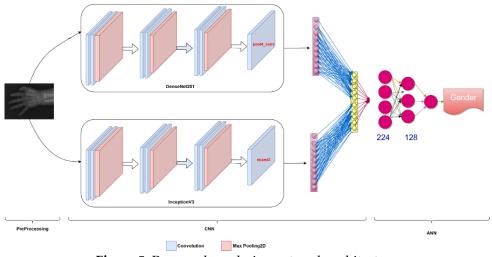
Figure 5. Proposed gendering network architecture

4.4. Performance Metrics

The proposed approach is evaluated using criteria such as accuracy, precision, sensitivity, and F-measure. The Recall value is the general name given to the metric value that shows how many of the transactions we need to predict positively. The accuracy is calculated by the ratio of the correctly predicted gender to the total data set. Precision is the general name of the metric value that allows us to test how many of the values we predicted as positive are positive. The F-measure is the harmonic mean of recall and precision. F-measures reveal realistic results that include all error costs in unevenly distributed datasets or datasets with unknown distribution. Equations of metrics are given in Table 3.

Table 3. Mathematical formulas of the performance parameters.

Parameter	Formula
Accuracy	(TP + TN)/(TP + TN + FP + FN)*100
Recall	TP/(TP + FN)
Precision	TP/(TP + FP)
F-Measure	{2 × (Recall × Precision)}/(Recall + Precision)

In these equations, T, F, P, and N stand for True, False, Positive, and Negative, respectively. For instance, TP represents the number of correctly classified positive samples, and FN represents the number of incorrectly classified negative samples.
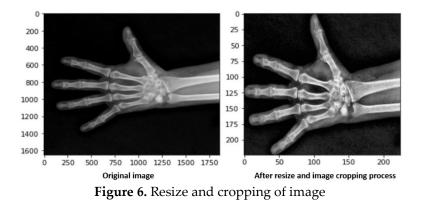
5. EXPERIMENTAL RESULTS

It has been determined that the sizes of the images obtained for this study are different. Inspired by the previous work of Ozdemir et al. [10], the image dimensions were set to 500*500 to prevent the loss of valuable information in the image. A special data generator method (imagedatagenerator) has been applied to keep and process images at this scale in RAM without any problems. In this method, 16 images with 500*500 dimensions were taken and processed in each step. Since ram error is received when more images are taken, all models are also taken as batch_size 16. Using this unique data generator method, the number of images up to batch size is given to the system in each epoch instead of all images. Thus, the encountered RAM problem has been eliminated.

To increase the models' performance, an automatic cropping process has been applied to the images. The state of an image after resizes and crop operation is given in figure 6.

Data augmentation was applied to the images as they were taken from various angles. The

normalization method of the transfer learning model used, rotation between -15 and +15, and 5% horizontal and vertical shift operations were applied to the training dataset during data augmentation. Only normalization method was applied to the images kept for validation and testing datasets. Learning rate 0.0001 epoch 100, batch size 12 and softmax as activation function were used.

Figure 6. Resize and cropping of image

Different metrics were used to evaluate the performance of the models. Of the total 13935 images in the dataset, 10069 (70%) were used for training, 1776 (15%) for validation and 2090 (15%) for testing. The performance values obtained for the test dataset of the 4 models are presented in Table 4.

Table 4. Results according to transfer learning models

Method	Accuracy	Recall	Precision	F1-Measure
InceptionV3	**94.71%**	**94.71%**	**94.72%**	**94.71%**
ResNet152	92.93%	92.92%	92.93%	94.91%
DenseNet201	94.23%	94.23%	94.35%	94.24%

Table 4 shows successful results were obtained with DenseNet201 and InceptionV3 models. Hence, to improve the model success, a hybrid model consisting of these two models was created. For this purpose, 2 ways are preferred. First, experimental studies were carried out by taking the layers of the models by trial and error. One of these experimental studies; A new hybrid model was created by taking the "mixed1" layer from the InceptionV3 model and the "Pool2_pool" layer from the DenseNet201 model. This model is called "Mixed1Pool2" model. As a result of this and similar experimental studies, it has been determined that this method gives good results, as seen in Table 5. As an alternative, a method that decides by looking at the effect of the layers of the trained model on the images was used and named as cross-layer display method. In this method, the layers were decided by applying the interlayer imaging method on both models and a new hybrid model was created. In the cross-layer display method; It was decided to select the layers by visualizing which features were learned by each layer in which the model processed the input image and which features passed to the deeper layers. Cross-layer display method shows which layer of the models learning took place and from which layer learning did not occur. Figure 7 shows visualized effects of InceptionV3 model layers on an image a with this method.
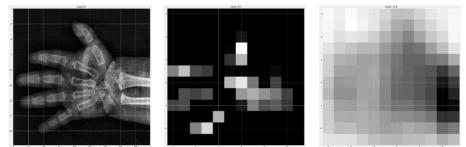

Figure 7. The effect of layers on the image

The IncepitonV3 model consists of 317 layers. In Figure 7, a normal image is given on the left. The second image shows the original image at Layer 97 of the IncepitonV3 model. The last image shows the image in Layer 134 of the IncepitonV3 model. As can be seen in Figure 7, in a model consisting of 317 layers, visuals can be understood up to the 97th layer. Therefore, the effect of the layers after this point on the model performance is very limited. The model learning takes place up to the "Mixed3" layer in the InceptionV3 model, and learning takes place up to the "pool4_conv" layer in the DenseNet201 model. Therefore, a new hybrid model named "Mixed3Pool4" was created by taking the relevant layers from each models. Table 5 shows the performance of hybrid model.

Table 5. Results according to transfer learning models

Method	Accuracy	Recall	Precision	F1-Measure
Mixed1Pool2	95.19%	95.19%	95.26%	95.2%
Mixed3Pool4	**96.29%**	**96.29%**	**96.29%**	**96.28%**

The hybrid models give more successful results than traditional transfer learning models which is shown in Table 5. Among all the experimental studies, the highest scoring classification results were obtained with the Mixed3Pool4 model with an accuracy rate of 96.29%. Figure 8 shows the accuracy and error values of the Mixed3Pool4 model during training.
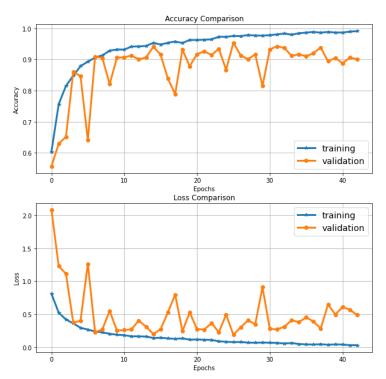

Figure 8. Training-validation accuracy and loss of Mixed3Pool4

For the Mixed3Pool4 model, the actual gender and predicted gender information of the images of a few randomly selected samples are given in Figure 9.

Figure 9. A sample of actual and predicted gender results obtained with the Mixed3Pool4 method

The Grad-CAM (Gradient weighted Class Activation Mapping) method is used to show the critical regions that the model focuses on while estimating the image of the trained model. Figure 10 shows the heat map of the regions where the trained Mixed3Pool4 model focuses on images of males, Figure 11 on females.
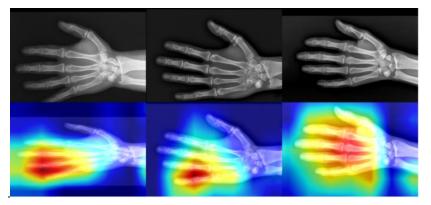

Figure 10. Heatmap (GradCam) on Male images of the trained Mixed3Pool4 model
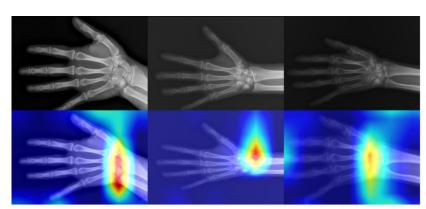
Figure 11. Heatmap (GradCam) of the trained Mixed3Pool4 model on the Female images

When we look at the heat maps, the Mixed3Pool4 model, in which we obtained the best results, focused mainly on the phalanx, carpal, and metacarpal bones in male and mainly on the distal radius and ulna in female. However, our model did not focus on the same points by gender in all images. Sometimes it focused on the carpal area for male, sometimes on the points between the fingers for female.

Images were divided into genders and decades of 10 to measure the success of the Mixed3Pool4 model. Then, with the Mixed3Pool4 model, the prediction success of the model was measured on the images divided into age groups according to gender. The success of the Mixed3Pool4 model on age groups according to gender in all images is shown in Table 6.

Table 6. Accuracies of the Mixed3Pool4 model by age groups and gender

Age Ranges	Male Success Rate (#Images)	Female Success Rate (#Images)	#Tested Image
2-10	0,978 (851)	0,969 (866)	1717
11-20	0,982 (1657)	0,990 (2512)	4169
21-30	0,986 (639)	0,990 (1372)	2011
31-40	0,992 (1196)	0,994 (802)	1998
41-50	0,996 (757)	0,988(755)	1512
51-60	0,988 (761)	0,977 (564)	1325
61-70	%100 (450)	0,975 (325)	775
71-79	0,996 (274)	0,910 (154)	428

As shown in Table 6, the images are concentrated in the 10-30 age range. The images in other age groups are close to each other in terms of gender. Mixed3Pool4 model was more successful for male in the 2-10 age range. It has been determined that the model is more successful for females in the 10-40 age range and for male over 40 years.

In the literature, many studies have been carried out for gender estimation. In the studies, the success of the models was generally evaluated based on accuracy. Some of the studies carried out are given in Table 6. When the results of the proposed approach of this study and the studies in Table 6 are compared, it is seen that successful results are obtained with the proposed approach.

Table 7. Studies in the literature in which gender identification was made from hand-wrist images

Authors	Method	DataSet	Accuracy (%)
Karadayi et. al. [42]	Manuel	255 males and 155 females.	86.6%
Jee et. al. [43]	t-test + Discriminant analysis	167 Male ve 154 Female	Male=88.6% Female= 89.6%
Prabhat et. al. [44]	t-test + Discriminant analysis	15 males and 15 females	Male 80.0 % Female 86.7%
Yune et. al. [31]	CNN	5459 Female and 5148 Male	95.9%
DeSilva et. al. [45]	Manuel	150 adult males and 150 adult females	85.3%
Barnes et. al. [32]	Statistical	50 Male, 50 Female	91.8%
Proposed Method	**CNN**	**Male 6191, Female 7744**	**96.3%**

Many different methods were used for gender estimation using wrist length, and there were very few studies on gender estimation from wrist radiographs. As can be seen from Table 7, our study is the research with the highest number of images and achieved the highest success when looking at the studies in the literature.

6. CONCLUSION

The intrinsic diversity in hand-wrist morphologies among individuals, stemming from the multifaceted influence of environmental factors on human development, is an established phenomenon. Through the lens of this investigation, a pioneering decision support system has been meticulously devised to proffer rapid gender delineation, lending its potential to both legal contexts and the augmentation of model expert analyses. Gender identification, bearing substantial pertinence within the domain of forensic sciences, has garnered extensive attention through diverse scholarly inquiries endeavoring to discern gender via anatomical measurements.

This study, distinguished by its novel approach, embarks upon the avenue of gender estimation hinged on X-ray depictions of hand-wrist bone structures. Drawing from an expansive dataset culled from four distinct medical institutions, encompassing a demographic spanning ages 1 to 85 years, the repository amalgamates a total of 13,935 X-ray images. Ingeniously employing deep learning models, the exploratory phase unfolded through experimental forays into the potent DenseNet201 and InceptionV3 models. The culmination of this endeavor surfaced in the formulation of a pioneering hybrid model, a harmonious amalgamation of these aforementioned architectures.

Central to the model's design was the methodical determination of layers through visualization of transfer learning algorithmic strata. Employing this strategy, layers extending up to the threshold of learning saturation were strategically embraced. The culmination of these efforts yielded a compelling success rate of 96.3% for the proposed hybrid model. This pivotal study stands as the inaugural investigation within our nation's confines, engendering gender estimation through deep learning constructs, predicated upon an exclusive repository of hand-wrist radiographs.

Additionally, this study ventures into the intriguing landscape of age-dependent gender distinctions, ascertaining nuanced variances within different age cohorts. The findings underscore the superiority of gender identification accuracy for boys within the 2-10 age bracket, with a converse trend in the 10-40 range, favoring females. Remarkably, this pattern inverts once more within the 40-79 age range, favoring males. The discernment that females manifest heightened accuracy prior to the age of 40, contrasted by males emerging triumphant after 40, conceivably aligns with the specter of osteoporosis. As a prevalent ailment, osteoporosis [46] marks a poignant disparity in bone density and microarchitecture, thus presenting significant public health challenges. Delving further, the differential

impact of osteoporosis on genders emerges, wherein the protective role of premenopausal estrogen grants females comparative resilience, in contrast to the heightened susceptibility post-menopause [47].

this inquiry yields a resounding success in gender identification through X-ray imagery, underpinned by the innovative hybrid fusion of DenseNet and InceptionV3 deep learning paradigms. The emergent ramifications of this study crystallize in a cascade of opportunities for future exploration. The rich tapestry of insights unveiled within these findings augurs the cultivation of novel paradigms in forensic gender identification, potentially fostering interdisciplinary collaborations and spurring innovation within medical imaging and artificial intelligence realms alike.

Declaration of Competing Interest

The authors declare no conflict of interest

Funding

The author(s) received no financial support for this article's research, authorship, and/or publication.

Declaration of Ethical Standards

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

REFERENCES

[1]   M. Y. İşcan ve M. Steyn, "The Human Skeleton in Forensic Medicine," 3rd ed.

[2]   A. P. Agelarakis, "Studies in Crime: An Introduction to Forensic Archaeology. By John R. Hunter, Charlotte A. Roberts, and Anthony L. Martin. Making Faces: Using Forensic and Archaeological Evidence. By John Prag and Richard Neave," Am J Archaeol, vol. 102, 1998. [Online]. Available: https://doi.org/10.2307/506107

[3]   M. Didi AL, R. R. Azman, ve M. Nazri, "Sex determination from carpal bone volumes: A Multi Detector Computed Tomography (MDCT) study in a Malaysian population," Leg Med, vol. 20, 2016. [Online]. Available: https://doi.org/10.1016/j.legalmed.2016.04.002

[4]   ILO Barrier ve EN L'Abbé, "Sex determination from the radius and ulna in a modern South African sample," Forensic Sci Int, vol. 179, 2008. [Online]. Available: https://doi.org/10.1016/j.forsciint.2008.04.012

[5]   T. K. Black, "A new method for assessing the sex of fragmentary skeletal remains: Femoral shaft circumference," Am J Phys Anthropol, vol. 48, 1978. [Online]. Available: https://doi.org/10.1002/ajpa.1330480217

[6]   L. R. Frutos, "Metric determination of sex from the humerus in a Guatemalan forensic sample," Forensic Sci Int, vol. 147, 2005. [Online]. Available: https://doi.org/10.1016/j.forsciint.2004.09.077

[7]   M. Y. Işcan, "Forensic anthropology of sex and body size," Forensic Sci Int, vol. 147, 2005. [Online]. Available: https://doi.org/10.1016/j.forsciint.2004.09.069

[8]   K. Sakaue, "Sexual determination of long bones in recent Japanese," Anthropol Sci, vol. 112, 2004. [Online]. Available: https://doi.org/10.1537/ase.00067

[9]   M. Šlaus, Ž. Bedić, D. Strinović, ve V. Petrovečki, "Sex determination by discriminant function analysis of the tibia for contemporary Croats," Forensic Sci Int, vol. 226, 2013. [Online]. Available: https://doi.org/10.1016/j.forsciint.2013.01.025

[10]   C. Ozdemir, M. A. Gedik, ve Y. Kaya, "Age Estimation from Left-Hand Radiographs with Deep Learning Methods," Trait du Signal, vol. 38, 2021. [Online]. Available: https://doi.org/10.18280/ts.380601

[11]   S. Gornale, A. Patil, ve C. V, "Fingerprint based Gender Identification using Discrete Wavelet Transform and Gabor Filters," Int J Comput Appl, vol. 152, 2016. [Online]. Available: https://doi.org/10.5120/ijca2016911794

[12]   X. Li, X. Zhao, Y. Fu, ve Y. Liu, "Bimodal gender recognition from face and fingerprint," in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition

[13]   R. K, A. Patil, ve S. Gornale, "Fusion of Features and Synthesis Classifiers for Gender Classification using Fingerprints," Int J Comput Sci Eng, vol. 7, 2019. [Online]. Available: https://doi.org/10.26438/ijcse/v7i5.526533

[14]   S. Aryanmehr, M. Karimi, ve F. Z. Boroujeni, "CVBL IRIS Gender Classification Database Image Processing and Biometric Research, Computer Vision and Biometric Laboratory (CVBL)," In: 2018 3rd IEEE International Conference on Image, Vision and Computing, ICIVC 2018.

[15]   M. Singh, S. Nagpal, M. Vatsa, ve diğerleri, "Gender and ethnicity classification of Iris images using deep class-encoder," In: IEEE International Joint Conference on Biometrics, IJCB 2017, 2018.

[16]   V. Thomas, N. V. Chawla, K. W. Bowyer, ve P. J. Flynn, "Learning to predict gender from iris images," In: IEEE Conference on Biometrics: Theory, Applications and Systems, BTAS'07, 2007.

[17]   S. S. Gornale, A. Patil, M. Hangarge, ve R. Pardesi, "Automatic human gender identification using palmprint," In: Smart Computational Strategies: Theoretical and Practical Aspects, 2019.

[18]   E. Tartaglione, C. A. Barbano, C. Berzovini, ve diğerleri, "Unveiling COVID-19 from chest x-ray with deep learning: A hurdles race with small data," Int J Environ Res Public Health, vol. 17, 2020. [Online]. Available: https://doi.org/10.3390/ijerph17186933

[19]   Y. Kaya ve Ö. F. Ertuğrul, "Gender classification from facial images using gray relational analysis with novel local binary pattern descriptors," Signal, Image Video Process, vol. 11, 2017. [Online]. Available: https://doi.org/10.1007/s11760-016-1021-3

[20]   L. Bui, D. Tran, X. Huang, ve G. Chetty, "Classification of gender and face based on gradient faces," In: 3rd European Workshop on Visual Information Processing, EUVIP 2011 - Final Program, 2011.

[21]   A. Hasnat, S. Halder, D. Bhattacharjee, ve M. Nasipuri, "A proposed system for gender classification using the lower part of the face image," In: Proceedings - IEEE International Conference on Information Processing, ICIP 2015, 2016.

[22]   M. Wu ve Y. Yuan, "Gender classification based on geometry features of palm image," Sci World J, vol. 2014. [Online]. Available: https://doi.org/10.1155/2014/734564

[23]   M. V. Rajee ve C. Mythili, "Gender classification on digital dental x-ray images using deep convolutional neural network," Biomed Signal Process Control, vol. 69, 2021. [Online]. Available: https://doi.org/10.1016/j.bspc.2021.102939

[24]   S. Babacan, S. Isiklar, I. M. Kafa, ve G. Gokalp, "Redesign of missing mandible by determining age group and gender from morphometric features of the skull for facial reconstruction (approximation)," Archaeol Anthropol Sci, vol. 13, 2021. [Online]. Available: https://doi.org/10.1007/s12520-021-01315-2

[25]   T. Danisman, I. M. Bilasco, ve J. Martinet, "Boosting gender recognition performance with a fuzzy inference system," Expert Syst Appl, vol. 42, 2015. [Online]. Available: https://doi.org/10.1016/j.eswa.2014.11.023

[26]   P. Gnanasivam ve S. Muttan, "Gender classification using ear biometrics," In: Lecture Notes in Electrical Engineering, 2013.

[27]   R. Djemili, H. Bourouba, ve M. C. A. Korba, "A speech signal-based gender identification system using four classifiers," In: Proceedings of 2012 International Conference on Multimedia Computing and Systems, ICMCS 2012, 2012.

[28]   L. Zimman, "Transgender voices: Insights on identity, embodiment, and the gender of the voice," Lang Linguist Compass, vol. 12, 2018. [Online]. Available:

https://doi.org/10.1111/lnc3.12284

[29]   A. B. A. Graf ve F. A. Wichmann, "Gender classification of human faces," In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2002.

[30]   X. Li, S. J. Maybank, S. Yan, ve diğerleri, "Gait components and their application to gender recognition," IEEE Trans. Syst. Man Cybern. Part C Appl. Rev., vol. 38, 2008.

[31]   S. Yune, H. Lee, M. Kim, ve diğerleri, "Beyond Human Perception: Sexual Dimorphism in Hand and Wrist Radiographs Is Discernible by a Deep Learning Model," J Digit Imaging, vol. 32, 2019. [Online]. Available: https://doi.org/10.1007/s10278-018-0148-x

[32]   A. E. Barnes, D. T. Case, S. E. Burnett, ve P. Mahakkanukrauh, "Sex estimation from the carpal bones in a Thai population," Aust J Forensic Sci, vol. 52, 2020. [Online]. Available: https://doi.org/10.1080/00450618.2019.1620856

[33]   N. Vila-Bianco, R. R. Vilas, M. J. Carreira, ve I. Tomás, "Towards deep learning reliable gender estimation from dental panoramic radiographs," In: CEUR Workshop Proceedings, 2020.

[34]   J. Bewes, A. Low, A. Morphett, ve diğerleri, "Artificial intelligence for sex determination of skeletal remains: Application of a deep learning artificial neural network to human skulls," J Forensic Leg Med, vol. 62, 2019. [Online]. Available: https://doi.org/10.1016/j.jflm.2019.01.004

[35]   E. Korot, N. Pontikos, X. Liu, ve diğerleri, "Predicting sex from retinal fundus photographs using automated deep learning," Sci Rep, vol. 11, 2021. [Online]. Available: https://doi.org/10.1038/s41598-021-89743-x

[36]   S. Gornale S, A. Patil, ve R. K, "Multimodal Biometrics Data Analysis for Gender Estimation Using Deep Learning," Int J Data Sci Anal, vol. 6, 2020. [Online]. Available: https://doi.org/10.11648/j.ijdsa.20200602.11

[37]   Ş. Kılıç ve Y. Doğan, "Deep learning-based gender identification using ear images," Traitement du Signal, vol. 40, 2023. [Online]. Available: https://doi.org/10.18280/ts.400431

[38]   Y. Doğan, "Derin Öğrenme Yöntemleriyle Çapraz Veri Seti Değerlendirmesi Altında COVID-19 Tespiti," Gazi University Journal of Science Part C: Design and Technology, vol. 11, no. 3, 2023.

[39]   M. A. Morid, A. Borjali, ve G. Del Fiol, "A scoping review of transfer learning research on medical image analysis using ImageNet," Comput. Biol. Med., vol. 128, 2021.

[40]   C. Szegedy, W. Liu, Y. Jia, ve diğerleri, "Going deeper with convolutions," In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[41]   G. Huang, Z. Liu, L. Van Der Maaten, ve K. Q. Weinberger, "Densely connected convolutional networks," In: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017.

[42]   B. Karadayi, A. Kaya, H. Koc, ve diğerleri, "Sex determination for using hand and wrist measurements in the Turkish population," Turkish J Forensic Med, vol. 28, 2014. [Online]. Available: https://doi.org/10.5505/adlitip.2014.98700

[43]   S. C. Jee, S. Bahn, ve M. H. Yun, "Determination of sex from various hand dimensions of Koreans," Forensic Sci Int, vol. 257, 2015. [Online]. Available: https://doi.org/10.1016/j.forsciint.2015.10.014

[44]   M. Prabhat, S. Rai, M. Kaur, ve diğerleri, "Computed tomography-based forensic gender determination by measuring the size and volume of the maxillary sinuses," J Forensic Dent Sci, vol. 8, 2016. [Online]. Available: https://doi.org/10.4103/0975-1475.176950

[45]   R. DeSilva, A. Flavel, ve D. Franklin, "Estimation of sex from the metric assessment of digital hand radiographs in a Western Australian population," Forensic Sci Int, vol. 244, 2014. [Online]. Available: https://doi.org/10.1016/j.forsciint.2014.08.019

[46]   T. Sozen, L. Ozisik, ve N. Calik Basaran, "An overview and management of osteoporosis," Eur J Rheumatol, vol. 4, 2017. [Online]. Available: https://doi.org/10.5152/eurjrheum.2016.048

[47]   F. Cosman, S. J. de Beur, M. S. LeBoff, ve diğerleri, "Clinician's Guide to Prevention and Treatment of Osteoporosis," Osteoporos Int, vol. 25, 2014. [Online]. Available:

https://doi.org/10.1007/s00198-014-2794-2

[48]   Ş. Öztürk, U. Özkaya, ve M. Barstuğan, "Classification of Coronavirus (COVID-19) from X-ray and CT images using shrunken features," Int J Imaging Syst Technol, vol. 31, no. 1, 2021. [Online]. Available: https://doi.org/10.1002/ima.22432