IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 9, 2021, accepted July 24, 2021, date of publication July 28, 2021, date of current version August 10, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3101040

Classification of Canine Maturity and Bone Fracture Time Based on X-Ray Images of Long Bones

GÜLNUR BEGÜM ERGÜN AND SELDA GÜNEY
Department of Electrical and Electronics Engineering, Başkent University, 06790 Ankara, Turkey
Corresponding author: Selda Güney (seldaguney@baskent.edu.tr)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Ankara Metropolitan Municipality Stray Animals Temporary Nursing Home.

ABSTRACT Veterinarians use X-rays for almost all examinations of clinical fractures to determine the appropriate treatment. Before treatment, vets need to know the date of the injury, type of the broken bone, and age of the dog. The maturity of the dog and the time of the fracture affects the approach to the fracture site, the surgical procedure and needed materials. This comprehensive study has three main goals: determining the maturity of the dogs (Task 1), dating fractures (Task 2), and finally, detecting fractures of the long bones in dogs (Task 3). The most popular deep neural networks are used: AlexNet, ResNet-50 and GoogLeNet. One of the most popular machine learning algorithms, support vector machines (SVM), is used for comparison. The performance of all sub-studies is evaluated using accuracy and F1 score. Each task has been successful with different network architecture. ResNet-50, AlexNet and GoogLeNet are the most successful algorithms for the three tasks, with F1 scores of 0.75, 0.80 and 0.88, respectively. Data augmentation is performed to make models more robust, and the F1 scores of the three tasks were 0.80, 0.81, and 0.89 using ResNet-50, which is the most successful model. This preliminary work can be developed into support tools for practicing veterinarians that will make a difference in the treatment of dogs with fractured bones. Considering the lack of work in this interdisciplinary field, this paper may lead to future studies.

INDEX TERMS Bone age, bone fractures, classification, convolutional neural networks, deep learning.

I. INTRODUCTION

In recent years, many fields of study, particularly biomedicine, have been positively affected by the phenomenon of deep learning [1]. Recent studies have examined disease and fracture detection, organ and tissue segmentation, and many more applications using magnetic resonance (MR) and tomography images with high success rates thanks to deep learning algorithms [2], [3].

However, these successful applications are valid for human medicine. Veterinary medicine is as important to the world economy as human medicine, and unfortunately, does not benefit from the blessings of engineering [4]. The importance of veterinary medicine can be emphasized with this common expression: "If human medicine is for people, veterinary medicine is for humanity."

The associate editor coordinating the review of this manuscript and approving it for publication was Carmelo Militello.

Like human medicine, veterinary medicine specializes in different fields, such as orthopedics, cardiology, urology, and virology. These fields required specialized education [5]. This study is aimed to help general practice veterinarians who do not have additional training in orthopedics or surgery.

A few interdisciplinary studies integrate engineering and veterinary medicine. A study on pigs can be given as an example of animal image processing [6]. The aim is to develop a fully automated pig skeleton segmentation method from computed tomography (CT) scanning. Applying a convolutional neural network (CNN), the researchers achieved 95% success and claimed that using deep learning methods increased robustness and reduced the need for manual control. According to the paper, the outcome of the image segmentation process can be used to improve pig breeding [6].

In another paper, McEvoy *et al.* aimed to classify images from canine pelvic radiographs using two machine learning

methods: partial least square discriminant analysis and artificial neural network (ANN). They used 256 images of 60 dogs (200 images for training and 56 images for the test set). The successes of these two methods are promising, and the results can be integrated into the veterinary field [7].

To show the potential of deep learning techniques in medicine, Vinicki *et al.* tried to predict the number of reticulocytes in microscope images of cat blood smears using a single-shot detector (SSD) and CNN architectures. They used 800 images and achieved 98.7% success [8].

Another work by Banzoto *et al.* aimed to predict whether a lesion is a meningioma or glioma on canine MR images using CNN. According to their work, classification accuracy is 94% [9].

Xiaoping Huang *et al.* also reviewed a detection model using faster R-CNN (region-based CNN) on cow tails. The study is useful for determining the body condition score (BCS) of cows because BCS is a parameter that states the muscle capacity of any cow [10].

Another paper reviewed CAD systems for breast cancer diagnosis. The authors claimed that there had been no published article about PET/MR for breast cancer diagnosis using machine learning methods yet [11].

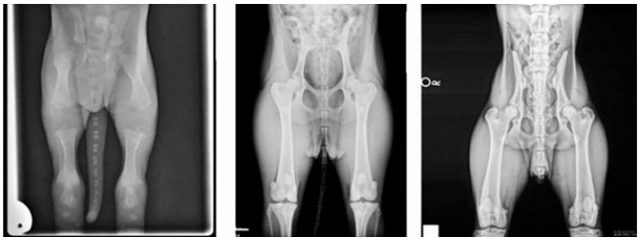
Arsomngern *et al.* investigated a lung lesion problem in pets using 2862 thoracic X-ray images taken from both dogs and cats. The results showed that CNN accurately identified the lesion in 79.6% of the cases [12].

The primary research purpose of our data set was classification of the types of long bones in dogs. To do that, 1819 images were divided into four classes according to the type of long bones: femur, humerus, radius-ulna and tibia. Using AlexNet, a classification accuracy of 0.82 was achieved. The result of that study was promising for future work [13].

Much has been written about deep learning applications, mostly for human medicine [14]–[16]. Our study is quite different and significant for the field. Because few similar studies have investigated animal health with engineering solutions in detail, outcomes from this study can be used as an assistant to veterinarians. Veterinarians will be prepared before the surgery and will be able to decide how the surgery will be performed. Additionally, several misinterpretations made in some special cases can be prevented by this guide. Our work is comprehensive in terms of the number of patients, radiograph images, and the perspective on the problems of animals.

The study is divided into three parts. The first task aims to determine the maturity of the dog by looking at the X-ray images. This is a challenging problem because all the dogs in the shelter are stray animals whose exact age cannot be known.

Veterinarians can predict the maturity of an animal using radiographs. In most mammals, including dogs, the growth plates (epiphyseal plates) at the bone edges close after sexual maturity, and then bone growth stops [17]. Figures 1 and 2 compare young and adult dog radiographs and growth plates.



(a) (b) (c)

FIGURE 1. Canine radiographs according to their age [18]. From the left; a) 2-week-old puppy, b) 7-week-old puppy, and c) 1-year-old adult dog.

(a) (b)

FIGURE 2. The puppy and adult dog growth plates [19]. From the left; a) puppy (growth plate is still open), b) adult (growth plate is closed).

The growth plate in the puppy is still open and more obvious than the adult dog.

Many studies based on the human system are related to predicting bone age [20]–[24]. Conversely, for animals, no similar research using deep learning methods exists to the best of our knowledge. This uninvestigated issue can be important due to the surgical procedure. For instance, anesthesia dosage is adjusted by considering the age and weight of the animal. Incorrect doses may cause serious problems [25].

The second task of this study is to determine when a fracture occurred. This information is important because it indicates how difficult surgery to repair the fracture will be. In addition, within limited resources, surgical priority can be determined among stray dogs, many of which suffer from similar orthopedic problems. Old fractures that have already begun to heal in the wrong position, called malunion, are more difficult to treat [26].

This area has not been explored in academic studies. Generally, humans know when they were injured. However, in most cases involving stray animals, veterinarians do not know when the fracture occurred. Therefore, this area of research can be useful.

Some papers state that callus formation begins 5–10 days after the fracture [27]. This study used a threshold time of 1 week. The fracture was categorized as new if it had occurred within 1 week, or old otherwise. Figure 3 demonstrates the bones' healing process in detail.

X-ray images that distinguish callus formation can be seen in Figure 4. As the healing of the fracture continues, the callus formation becomes larger and larger [28], [29].

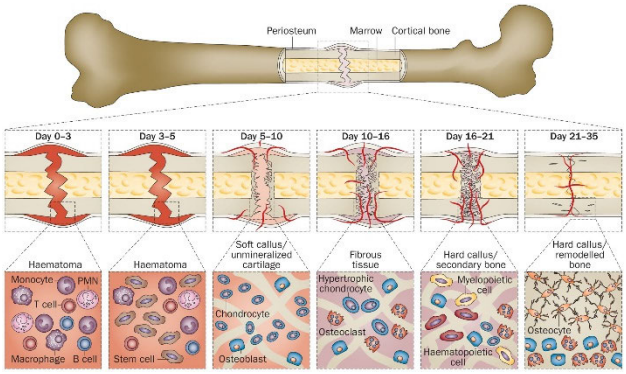


FIGURE 3. Healing process of a bone [27].

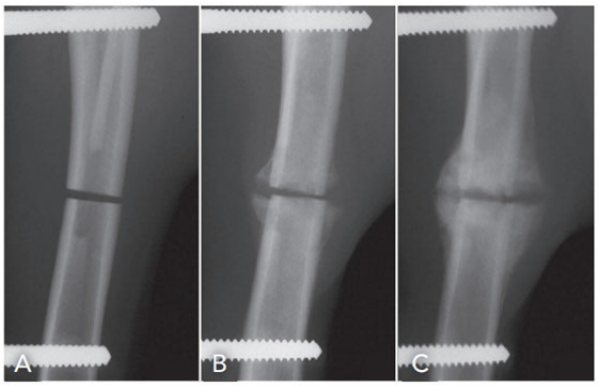


FIGURE 4. Healing process of a bone from X-ray images [28]. A) Newly formed fracture, B) Initial phase of tissue healing, C) Larger callus formation.

The final task of our study is detecting fractures. Detection of bone fractures is one of the most studied subjects in orthopedics medicine [30]–[38]. Many of these research studies apply transfer learning as a reasonable way to detect fractures. It leads us to apply the same methods for animals.

One study with a wide perspective on bone fractures was done by Abbas *et al.* Their goal is to detect fractures and recognize the fracture category. Despite using a small, human-based data set, the overall accuracy of their proposed method is 97% [39].

One of the latest studies on cats and dogs was done by Zhou *et al.* They presented an intelligent image interpretation system. The system can classify between X-ray images with and without lesions first. Then the system can localize the lesions if the images are classified as with lesions. They used the human musculoskeletal radiograph (MURA) data set and 500 X-ray images from 141 animals (123 dogs, 18 cats). They proposed two different ways for training process. The overall accuracies of their proposed methods are 76.6% and 77.4% [40].

The contributions of our study can be summarized as follows:

1. Creation of a comprehensive data set. Each X-ray image in the data set has been carefully labeled by a specialist orthopedist veterinarian. The data set is available online for other researchers. Therefore, the study can be an asset, opening the door to novel research.

2. The groundwork has been completed for the development of supporting tools for practicing veterinarians. The health problems of animals have not received as much attention as human health problems. This work begins to close the gap.

3. The effect of the data augmentation technique on canine X-ray images is investigated. To the best knowledge of the authors, previous studies have not examined this category of images.

II. MATERIAL AND METHODS

A. DATA SET

The success of any deep learning algorithm depends on many parameters, and the key factor of the success is data. The quality, number and variety of data have crucial importance for accuracy. The X-ray images in the data set were obtained from many dogs belong to Ankara Metropolitan Municipality Stray Animals Temporary Nursing Home. The labeling process was carried out by a specialist orthopedist veterinarian.

The X-ray images were taken by a Fujifilm CR-IR-392 and visualized by the FVS-100 interface program. During the visualization process, they were turned into *.png extension images. In this way, 3212 images were examined, and 1120 images were eliminated because of noise. Figure 5 shows examples of X-ray images of long bones in the data set.

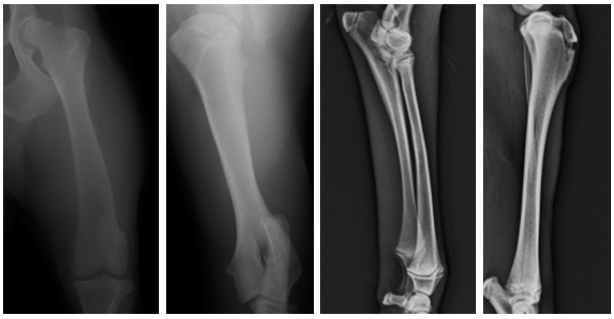


FIGURE 5. Examples of long bones in the data set. From the left; a) femur, b) humerus, c) radius-ulna, and d) tibia.

For better visualization, the bones in dogs are given in Figure 6, and the names of the long bones are shown in a rectangle. Although the fibula is a long bone, it was not included in this study because the fibula does not have any effective role in carrying body weight [41].

B. DATA AUGMENTATION

Data augmentation was performed to ensure the robustness of the models. Many data augmentation methods exist in the literature; one of the most basic methods, changing the brightness of the image, is preferred because of its ease of application [42]. The gamma value is taken as its default

value of 1, and the range of values in the output images is reversed. The number of images allocated for different tasks was doubled with this augmentation method.

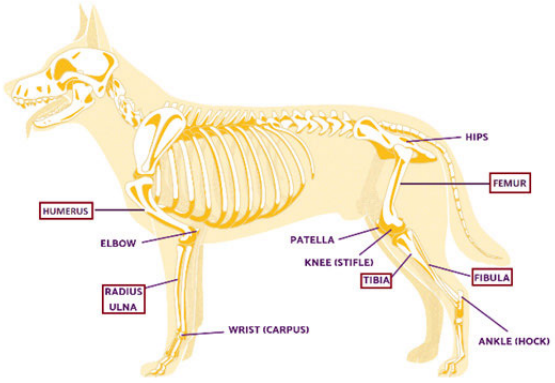


FIGURE 6. Main bones of a canine (dog) [43].

C. METHODS

CNN architectures were used in this study because of their success in image processing [44]. A convolutional neural network consists of five fundamental layers: An input layer, convolution layers, pooling layers, a fully connected layer, and an output layer [45]. The purpose of the convolution layer is to extract features from the input image by performing a dot product between image and filter matrices. After the convolution layer, the pooling layer is generally used to reduce the dimensions of the featured matrix. Finally, the output from the network matrix is flattened and ready for the classification process in the fully connected layer.

In all sub-studies in this paper, training and test sets were randomly selected as 80% and 20% of the large data set, respectively. The dimensions of all input images are 200 $\times$ 200 $\times$ 3 pixels. By applying the raw data to the CNN model, the features in the image were obtained and then classified with multi-class support vector machines (SVM) [46]. AlexNet, GoogLeNet and ResNet-50 architectures were used as transfer learning methods [47]–[49]. SVM was used for feature extraction and classification without any deep network because SVM has been proven its success with high dimensional data sets [50]. In the SVM algorithm, feature point locations were selected using the grid method with an 8 $\times$ 8 grid step. The number of image feature points was reduced for each task by keeping 80 percent of the strongest features from each category. After extracting features, the SVM classifier was trained on the training set and finally evaluated the performance of the classifier using the test set.

The performance of all sub-studies was evaluated using confusion matrices. The main performance metrics of the studies are accuracy and F1 score.

The confusion matrix represents the performance of a classifier. Several metrics are commonly used from confusion matrices to show the accuracy of the classifier. In this work, the accuracies of all models are calculated using (1).

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FN} + \text{FP}) \quad (1)$$

where the true negative (TN) parameter shows the number of negative examples classified accurately. Similarly, true positive (TP) indicates the number of positive examples classified accurately. False positive (FP) means the number of actual negative examples classified as positive; false negative (FN) is the number of actual positive examples classified as negative.

Accuracy is not the only metric for measuring the performance of the model. Other useful metrics for representing performance include sensitivity and precision. Precision is a metric in cases where FP is a greater concern than FN. Sensitivity is a useful metric in cases where FN is a greater concern than FP. It is not possible to say whether precision or sensitivity is more important. A combined metric, named the F1 score, can be calculated using (2).

$$\text{F1 score} = 2/((1/\text{sensitivity}) + (1/\text{precision})) \quad (2)$$

III. DISCUSSION AND EXPERIMENTAL RESULTS

A total of 1000 X-ray images were used for the first task. Half of these X-ray images belong to puppies, and the rest of them belong to adult dogs. The threshold between the two groups is 1 year. A specialist orthopedist veterinarian categorized the radiographs into the age groups without examining the dogs, so some inaccuracies are likely. Breed, sex, and other parameters were not considered for all sub-studies in this paper. All 1000 images featured long bones.

A total of 410 X-ray images were used for the second task of dating the fractures. Of these images, 304 are new fractures, and 106 are old, which means that an observable healing process has started. The threshold time for this task is 1 week. As with Task 1, a veterinary orthopedist examined and labeled the images.

Finally, a total of 2027 X-ray images were used for Task 3, including 481 images randomly selected from 1548 intact and 479 broken bones. After applying all the algorithms, classification accuracies for all the tasks are tabulated in Tables 1, 2 and 3.

Table 1 shows that ResNet-50 is the best choice, followed by GoogLeNet. Increasing the depth of the network appears to positively affect classification accuracy because ResNet-50, GoogLeNet and AlexNet consist of 50, 22 and 8 layers,

TABLE 1. Classification results for task 1: determining a dog's maturity.

Type of Network	*Classification Accuracy* (%)
AlexNet	63.5
GoogLeNet	74
ResNet-50	**76.5**
SVM	72.5

TABLE 2. Classification results for task 2: dating the fractures.

Type of Network	*Classification Accuracy (%)*
AlexNet	**80.95**
GoogLeNet	71.43
ResNet-50	73.81
SVM	64.29

TABLE 3. Classification results for task 3: fracture detection.

Type of Network	*Classification Accuracy (%)*
AlexNet	82.29
GoogLeNet	**88.54**
ResNet-50	85.42
SVM	73.07

TABLE 4. Classification results for task 1 with an augmented data set: determining a dog's maturity.

Type of Network	*Classification Accuracy (%)*
AlexNet	71
GoogLeNet	77.25
ResNet-50	**80**
SVM	65.25

TABLE 5. Classification results for task 2 with an augmented data set: dating fractures.

Type of Network	*Classification Accuracy (%)*
AlexNet	76.19
GoogLeNet	79.76
ResNet-50	**83.33**
SVM	59.52

TABLE 6. Classification results for task 3 with an augmented data set: fracture detection.

Type of Network	*Classification Accuracy (%)*
AlexNet	83.07
GoogLeNet	85.16
ResNet-50	**88.54**
SVM	68.26

respectively. This finding makes sense because the maturity of the animals is indicated by small details on the edge of the bone.

As the CNN architecture goes deeper, it reveals advanced features in images [51]. In this study, solutions are suggested for different recognition problems, and there is no best classification method for all the tasks. For instance, when we look at the second task, the depth of the network is inversely proportional to the accuracy. The AlexNet is the shallowest network among the ones studied.

Data augmentation is used to make the data set more robust. The number of images for each task was doubled. Afterward, the same procedures were performed on the augmented data set, and Tables 4, 5 and 6 were created. To better examine the classification performances, the confusion matrices are given in Table 7 and 8.

Table 7 shows that the old fractures can be detected better than new fractures due to the distinctive characteristic of the callus information around the fracture site. GoogLeNet has an 88.54% classification success in detecting fractures.

Transfer learning is indispensable when training large data sets like images in less time and achieve higher success. The SVM algorithm has not performed as well as the deep neural networks despite its success in other studies [52].

Without any image enhancement processes, the F1 scores of all the studies are promising. All F1 scores are above 0.66, and the highest score was obtained in Task 3. It is obvious that an X-ray of a broken bone has stronger distinguishing features than an X-ray of an adolescent-dog bone; thus, it is easier to classify.

False negatives have a higher concern than false positives in medical cases. Consequences can be more severe if actual positive cases are missed. Hence, sensitivity is the more important metric. Although the accuracy of SVM is low, the sensitivity rates are good. For this reason, SVM deserves consideration in future works.

Another investigation of the study is the effect of data augmentation on classification accuracy. In most cases, accuracies are better after the augmentation process. However, one point is notable: the decreasing effect on the accuracy of the SVM algorithm. It appears that SVM failed while working with a larger data set. Previous investigations have noted this phenomenon [53]. They claimed that the SVM algorithm performs better with higher quality samples than with a higher quantity of samples. Our study results confirm this claim.

Compared to the transfer learning method, even the training process with the normal data set takes a long time; the training process with the augmented data set takes much more time. Therefore, SVM is not recommended for the augmented data set. Further, after the augmentation process, the model constructed using ResNet-50 outperformed the other models for all tasks. The data augmentation process positively improved the learning capability of this model.

TABLE 7. Confusion matrices for all tasks.

Type of Network		*Confusion Matrix for Task 1*		*Confusion Matrix for Task 2*		*Confusion Matrix for Task 3*	
		Predicted Values					
AlexNet	*Actual Values*	55	45	18	3	82	14
		28	72	5	16	20	76
		***Sensitivity* 0.72**	***Precision* 0.61**	***Sensitivity* 0.76**	***Precision* 0.84**	***Sensitivity* 0.79**	***Precision* 0.84**
		F1 Score: 0.66		***F1 Score: 0.80***		***F1 Score: 0.81***	
GoogLeNet		79	21	17	4	85	11
		31	69	8	13	11	85
		***Sensitivity* 0.69**	***Precision* 0.76**	***Sensitivity* 0.62**	***Precision* 0.76**	***Sensitivity* 0.88**	***Precision* 0.88**
		F1 Score: 0.72		***F1 Score: 0.68***		***F1 Score: 0.88***	
ResNet-50		82	18	18	3	80	16
		29	*71*	*8*	13	12	84
		Sensitivity* *0.71	***Precision* *0.79***	***Sensitivity* *0.62***	***Precision* 0.81**	***Sensitivity* 0.87**	***Precision* 0.84**
		F1 Score: 0.75		***F1 Score: 0.70***		***F1 Score: 0.85***	
SVM		69	31	11	10	70	26
		24	*76*	*5*	16	26	70
		Sensitivity* *0.76	***Precision* *0.71***	***Sensitivity* *0.76***	***Precision* *0.61***	***Sensitivity* *0.73***	***Precision* *0.73***
		F1 Score: 0.73		***F1 Score: 0.68***		***F1 Score: 0.73***	

TABLE 8. Confusion matrices for all tasks with an augmented data set.

Type of Network		*Confusion Matrix for Task 1*		*Confusion Matrix for Task 2*		*Confusion Matrix for Task 3*	
		Predicted Values					
AlexNet	*Actual Values*	127	73	34	8	166	26
		43	157	12	30	39	153
		***Sensitivity* 0.78**	***Precision* 0.68**	***Sensitivity* 0.71**	***Precision* 0.79**	***Sensitivity* 0.80**	***Precision* 0.85**
		F1 Score: 0.73		***F1 Score: 0.75***		***F1 Score: 0.82***	
GoogLeNet		150	50	36	6	151	41
		41	159	11	31	16	176
		***Sensitivity* 0.79**	***Precision* 0.76**	***Sensitivity* 0.74**	***Precision* 0.84**	***Sensitivity* 0.92**	***Precision* 0.81**
		F1 Score: 0.77		***F1 Score: 0.79***		***F1 Score: 0.86***	
ResNet-50		158	42	40	2	158	34
		38	*162*	*12*	30	10	182
		***Sensitivity* 0.81**	***Precision* 0.79**	***Sensitivity* 0.71**	***Precision* 0.94**	***Sensitivity* 0.95**	***Precision* 0.84**
		F1 Score: 0.80		***F1 Score: 0.81***		***F1 Score: 0.89***	
SVM		144	56	21	21	131	61
		84	*116*	*13*	29	60	132
		Sensitivity* *0.58	***Precision* *0.67***	***Sensitivity* *0.69***	***Precision* *0.58***	***Sensitivity* *0.69***	***Precision* *0.68***
		F1 Score: 0.62		***F1 Score: 0.63***		***F1 Score: 0.68***	

The study has several limitations. First, all X-ray images were obtained from stray dogs, and there is no way to know the exact age of the animals. Similarly, the facts of the fractures are unknown. Dating the injury is a challenging problem for data set labeling. Although the data set has been examined by a specialist orthopedic veterinarian, some errors are possible. Secondly, the exact number of dogs in the data set is unknown. The associate veterinarian believes images from more than 1200 different dogs are included in the data set. Finally, the results are obtained from images without any enhancement process. Additionally, the parameters of the networks were determined by the trial-and-error method. Therefore, the accuracies of the networks do not reflect the highest achievable values.

The accuracy of future work can be increased by tuning network parameters or realizing the image enhancement process. Different augmentation methods can be tried on the data set and considered separately for each task.

IV. CONCLUSION

Whereas there are many successful deep learning applications for human medicine, veterinary medicine has not been studied much in this regard. Thus, the goal is to contribute engineering solutions to veterinary medicine. This paper investigated three topics to achieve this goal: determining a dog's maturity, dating long bone fractures, and long bone fracture detection.

CNN architectures were used in all tasks: AlexNet, ResNet-50 and GoogLeNet. SVM was used for comparison. The three deep learning methods are not very different, except for AlexNet in Task 2. ResNet-50, AlexNet, and GoogLeNet stand out for classification success in three tasks, respectively. The percentages can increase even more after an optimization process.

The effect of data augmentation is also investigated in this study. After doubling the data set, the F1 scores of most deep learning algorithms were increased. In contrast, the accuracy of SVM was negatively affected after the augmentation process. The SVM algorithm performs better with better quality samples than with a high quantity of samples.

The results obtained in the study are similar to results from studies conducted with human-based data sets. This is an indication that the transfer learning method is suitable for this study. It saves time during the training phase. The results showed that this research topic could be handled like most research on human-based data sets.

In short, we believe that this comprehensive work will contribute to the field. Considering the lack of studies in this interdisciplinary field, this paper could lead to future studies. Our next objective is to classify the type of long bone fracture in dogs, which is another area ready for enlightenment.

ACKNOWLEDGMENT

The authors would like to thank their biggest supporter, the special veterinarian orthopedist Tahsin Gürkan Ergün. They would also like to thank Ankara Metropolitan Municipality Stray Animals Temporary Nursing Home for the data set.

Ethics Approval: All procedures involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted. This article does not contain any studies with human participants.

Informed Consent: Informed consent was obtained from Ankara Metropolitan Municipality Stray Animals Temporary Nursing Home.

REFERENCES

[1] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, pp. 197–387, Jun. 2014.
[2] R. Zemouri, N. Zerhouni, and D. Racoceanu, "Deep learning in the biomedical applications: Recent and future status," *Appl. Sci.*, vol. 9, no. 8, p. 1526, Apr. 2019.
[3] X. Du, S. Yin, R. Tang, Y. Liu, Y. Song, Y. Zhang, H. Liu, and S. Li, "Segmentation and visualization of left atrium through a unified deep learning framework," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 4, pp. 589–600, Apr. 2020.
[4] P. Eyre, "Engineering veterinary education," *J. Vet. Med. Educ.*, vol. 29, no. 4, pp. 195–200, Dec. 2002.
[5] J. Buzzeo, D. Robinson, and M. Williams, "The 2014 RCVS survey of the veterinary profession," Inst. Employment Stud., Brighton, U.K., Sep. 2014.
[6] J. Kvam, L. E. Gangsei, J. Kongsro, and A. H. Schistad Solberg, "The use of deep learning to automate the segmentation of the skeleton from CT volumes of pigs1," *Transl. Animal Sci.*, vol. 2, no. 3, pp. 324–335, Jul. 2018.
[7] F. J. McEvoy and J. M. Amigo, "Using machine learning to classify image features from canine pelvic radiographs: Evaluation of partial least squares discriminant analysis and artificial neural network models," *Veterinary Radiol. Ultrasound*, vol. 54, no. 2, pp. 122–126, Mar. 2013.
[8] K. Vinicki, P. Ferrari, M. Belic, and R. Turk, "Using convolutional neural networks for determining reticulocyte percentage in cats," in *Proc. 20th ESVCP-ECVCP Meeting*, Athens, Greece, 2018.
[9] T. Banzato, M. Bernardini, G. B. Cherubini, and A. Zotti, "A methodological approach for deep learning to distinguish between meningiomas and gliomas on canine MR-images," *BMC Vet. Res.*, vol. 14, no. 1, p. 317, Dec. 2018.
[10] X. Huang, X. Li, and Z. Hu, "Cow tail detection method for body condition score using faster R-CNN," in *Proc. IEEE Int. Conf. Unmanned Syst. Artif. Intell. (ICUSAI)*, Nov. 2019, pp. 347–351.
[11] J. R. Burt, N. Torosdagli, N. Khosravan, H. RaviPrakash, A. Mortazi, F. Tissavirasingham, S. Hussein, and U. Bagci, "Deep learning beyond cats and dogs: Recent advances in diagnosing breast cancer with deep neural networks," *Brit. J. Radiol.*, vol. 91, no. 1089, Sep. 2018, Art. no. 20170545, doi: 10.1259/bjr.20170545.
[12] P. Arsomngern, N. Numcharoenpinij, J. Piriyataravet, W. Teerapan, W. Hinthong, and P. Phunchongharn, "Computer-aided diagnosis for lung lesion in companion animals from X-ray images using deep learning techniques," in *Proc. IEEE 10th Int. Conf. Awareness Sci. Technol. (iCAST)*, Oct. 2019, pp. 1–6.
[13] G. B. Ergün, S. Güney, and T. G. Ergün, "Köpeklerdeki uzun kemiklerin evriiimsel sinir Ağları kullanılarak sınıflandırılması," *Fırat Üniversitesi Fen Bilimleri Dergisi*, vol. 33, no. 1, pp. 125–132, 2021.
[14] M. A.Mohammed, K. H. Abdulkareem, B. Garcia-Zapirain, S. A. Mostafa, M. S. Maashi, A. S. Al-Waisy, M. Ahmed Subhi, A. Awad Mutlag, and D.-N. Le, "A comprehensive investigation of machine learning feature extraction and classification methods for automated diagnosis of COVID-19 based on X-ray images," *Comput., Mater. Continua*, vol. 66, no. 3, pp. 3289–3310, 2021.
[15] M. J. Awan, M. M. Rahim, N. Salim, M. Mohammed, B. Garcia-Zapirain, and K. Abdulkareem, "Efficient detection of knee anterior cruciate ligament from magnetic resonance imaging using deep learning approach," *Diagnostics*, vol. 11, no. 1, p. 105, Jan. 2021.
[16] A. S. Al-Waisy, M. A. Mohammed, S. Al-Fahdawi, M. S. Maashi, B. Garcia-Zapirain, K. H. Abdulkareem, S. A. Mostafa, N. M. Kumar, and D. N. Le, "COVID-DeepNet: Hybrid multimodal deep learning system for improving COVID-19 pneumonia detection in chest X-ray images," *Comput., Mater. Continua*, vol. 67, no. 2, 2021, pp. 2409–2429, 2021.
[17] K. SH, G. Trudel, and H. Uhthoff, "Review of growth plate closure compared with age at sexual maturity and lifespan in laboratory animals," *J. Amer. Assoc. Lab. Animal Sci.*, vol. 41, no. 5, pp. 21–26, 2002.
[18] A. Pishner. (Jan. 28, 2021). *The Greatest Argument Against Early Sterilization is the Orthopedic Repercussions*. [Online]. Available: https://valork9academy.com/2020/03/25/when-should-i-spay-neuter-my-dog/
[19] G. George. *How Old Should my Dog be to Canicross*. Accessed: Jan. 28, 2021. [Online]. Available: http://blog.dogfit.co.uk/2019/08/how-old-should-my-dog-be-to-canicross/
[20] S. Wang, Y. Shen, C. Shi, P. Yin, Z. Wang, P. W.-H. Cheung, J. P. Y. Cheung, K. D.-K. Luk, and Y. Hu, "Skeletal maturity recognition using a fully automated system with convolutional neural networks," *IEEE Access*, vol. 6, pp. 29979–29993, 2018.
[21] D. B. Larson, M. C. Chen, M. P. Lungren, S. S. Halabi, N. V. Stence, and C. P. Langlotz, "Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs," *Radiology*, vol. 287, no. 1, pp. 313–322, Apr. 2018.

[22] S. Wang, X. Wang, Y. Shen, B. He, X. Zhao, P. W. Cheung, J. P. Y. Cheung, K. D. luk, and Y. Hu, "An ensemble-based densely-connected deep learning system for assessment of skeletal maturity," *IEEE Trans. Syst., Man, Cybern., Syst.*, early access, Jun. 23, 2020, doi: 10.1109/TSMC.2020.2997852.

[23] S. Mutasa, P. D. Chang, C. Ruzal-Shapiro, and R. Ayyala, "MABAL: A novel deep-learning architecture for machine-assisted bone age labeling," *J. Digit. Imag.*, vol. 31, no. 4, pp. 513–519, Aug. 2018.

[24] V. I. Iglovikov, A. Rakhlin, A. A. Kalinin, and A. A. Shvets, "Paediatric bone age assessment using deep convolutional neural networks," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA 2018, ML-CDS* (Lecture Notes in Computer Science), vol. 11045. Cham, Switzerland: Springer, 2018, pp. 300–308.

[25] R. Bednarski, K. Grimm, R. Harvey, M. V. Lukasik, W. S. Penn, B. Sargent, and K. S. AAHA, "Anesthesia guidelines for dogs and cats," *J. Amer. Animal Hospital Assoc.*, vol. 47, no. 6, pp. 377–385, Nov. 2011.

[26] J. D. Pogue, F. S. Viegas, M. R. Patterson, D. P. Peterson, K. D. Jenkins, D. T. Sweo, and A. J. Hokanson, "Effects of distal radius fracture malunion on wrist joint mechanics," *J. Hand Surg.*, vol. 15, no. 5, pp. 721–727, 1990.

[27] W. Wang and K. W. K. Yeung, "Bone grafts and biomaterials substitutes for bone defect repair: A review," *Bioactive Mater.*, vol. 2, no. 4, pp. 224–247, Dec. 2017.

[28] M. Kapler and D. Dycus, "A practitioner's guide to fracture management part 2: Selection of fixation technique & external coaptation," Today's Vet. Pract., Sep./Oct. 2015, pp. 23–30.

[29] T. A. Einhorn and L. C. Gerstenfeld, "Fracture healing: Mechanisms and interventions," *Nature Rev. Rheumatology*, vol. 11, no. 1, pp. 45–54, Jan. 2015.

[30] M. Adams, W. Chen, D. Holcdorf, M. W. McCusker, P. D. Howe, and F. Gaillard, "Computer vs human: Deep learning versus perceptual training for the detection of neck of femur fractures," *J. Med. Imag. Radiat. Oncol.*, vol. 63, no. 1, pp. 27–32, Feb. 2019.

[31] A. Brett, C. G. Miller, C. W. Hayes, J. Krasnow, T. Ozanian, K. Abrams, J. E. Block, and C. van Kuijk, "Development of a clinical workflow tool to enhance the detection of vertebral fractures: Accuracy and precision evaluation," *Spine*, vol. 34, no. 22, pp. 2437–2443, Oct. 2009.

[32] S. W. Chung, S. S. Han, J. W. Lee, K.-S. Oh, N. R. Kim, J. P. Yoon, J. Y. Kim, S. H. Moon, J. Kwon, H.-J. Lee, Y.-M. Noh, and Y. Kim, "Automated detection and classification of the proximal humerus fracture by using deep learning algorithm," *Acta Orthopaedica*, vol. 89, no. 4, pp. 468–473, Jul. 2018.

[33] A. Zbigniew Starosolski, J. Herman Kan, and A. Annapragada, "CNN-based detection of distal tibial fractures in radiographic images in the setting of open growth plates," *Proc. SPIE*, vol. 11314, Mar. 2020, Art. no. 113143M.

[34] C.-L. Chin, Y.-L. Lin, and Y.-C. Liu, "Various types fracture labeling in bone radiographs using modified AC-GAN," in *Proc. Int. Conf. Technol. Appl. Artif. Intell. (TAAI)*, Nov. 2019, pp. 1–6.

[35] A. Y. Yang, L. Cheng, M. Shimaponda-Nawa, and H. Zhu, "Long-bone fracture detection using artificial neural networks based on line features of X-ray images," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Xiamen, China, 2019, pp. 2595–2602, doi: 10.1109/SSCI44817.2019.9002664.

[36] S. Beyaz, Salih, K. Açıcı, E. Sümer, *Derin Öğrenme ve Genetik Algoritma Yaklaşımları Kullanılarak X-Ray Görüntülerinde Femur Boyun Kırığı Tespiti*. İzmir, Türkiye: Uluslararası Sağlıkta Yapay Zeka Kongresi, 2020.

[37] D. H. Kim and T. MacKinnon, "Artificial intelligence in fracture detection: Transfer learning from deep convolutional neural networks," *Clin. Radiol.*, vol. 73, no. 5, pp. 439–445, May 2018.

[38] Y. Chen, *Classification of Atypical Femur Fracture With Deep Neural Networks*. Stockholm, Sweden: KTH Univ., 2019.

[39] W. Abbas, M. S. Adnan, M. A. Javid, W. Ahmad, and F. Ali, "Analysis of tibia-fibula bone fracture using deep learning technique from X-ray images," *Int. J. Multiscale Comput. Eng.*, vol. 19, no. 1, pp. 25–39, 2021.

[40] S. Zhou, E. Ahn, M. Fulham, and J. Kim, "Intelligent interpretation of veterinary musculoskeletal X-rays trained with human thoracic limb X-rays," Res. Square, Durham, NC, USA, 2021, doi: 10.21203/rs.3.rs-598900/v1.

[41] D. H. Carpenter, Jr., and R. C. Cooper, "Mini review of canine stifle joint anatomy," *Anat Histol Embryol.*, vol. 29, no. 6, pp. 321–329, Dec. 2000, doi: 10.1046/j.1439-0264.2000.00289.x.

[42] P. Haeberli and D. Voorhies, "Image processing by linear interpolation and extrapolation," *IRIS Universe Mag.*, vol. 28, pp. 8–9, Aug. 1994.

[43] (Jan. 30, 2021). *Dog Leg Anatomy in Human Terms*. [Online]. Available: https://orthodog.com/article/dog-leg-anatomy/

[44] K. Guo, L. Sui, J. Qiu, S. Yao, S. Han, Y. Wang, and H. Yang, "Angel-eye: A complete design flow for mapping CNN onto customized hardware," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI (ISVLSI)*, Jul. 2016, pp. 24–29.

[45] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989.

[46] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[47] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[48] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[49] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[50] I. Syarif, A. Prugel-Bennett, and G. Wills, "SVM parameter optimization using grid search and genetic algorithm to improve classification performance," *TELKOMNIKA (Telecommun. Comput. Electron. Control)*, vol. 14, no. 4, p. 1502, Dec. 2016.

[51] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6.

[52] P. Thanh Noi and M. Kappas, "Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery," *Sensors*, vol. 18, no. 2, p. 18, Dec. 2017.

[53] H. Yu, J. Yang, and J. Han, "Classifying large data sets using SVMs with hierarchical clusters," in *Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2003, pp. 306–315.

GÜLNUR BEGÜM ERGÜN was born in Eskişehir, Turkey, in 1993. She received the B.S. and M.S. degrees in electrical and electronics engineering from Başkent University, Ankara, Turkey, in 2016 and 2018, respectively, where she is currently pursuing the Ph.D. degree in electrical and electronics engineering. Since 2016, she has been a Research Assistant in electrical and electronics engineering with Başkent University. Her research interests include image processing, deep learning, and signal processing.

SELDA GÜNEY was born in Erzurum, Turkey, in 1982. She received the B.S., M.S., and Ph.D. degrees in electrical and electronics engineering from Karadeniz Technical University, Trabzon, Turkey, in 2004, 2007, and 2013, respectively. From 2005 to 2013, she was a Research Assistant with Karadeniz Technical University. Since 2013, she has been an Assistant Professor with the Department of Electrical and Electronics Engineering, Başkent University. She is the author of more than 30 articles. Her research interests include signal processing, gas sensor, machine learning, pattern recognition, and artificial intelligence.