수의학석사 학위논문

Deep Learning Estimation of Age in Geriatric Dogs Using Thoracic Radiographs

흉부 방사선 사진을 활용한 인공지능 기반 노령견 나이 추정에 관한 연구

2023년 02월

서울대학교 대학원

수의학과 임상수의학 전공

정 지 원

Deep Learning Estimation of Age

in Geriatric Dogs

Using Thoracic Radiographs

지도교수 최 지 혜

이 논문을 수의학석사 학위논문으로 제출함

2022년 12월

서울대학교 대학원

수의학과 임상수의학 전공

정 지 원

정지원의 수의학석사 학위논문을 인준함

2023년 01월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

Abstract

Deep Learning Estimation of Age in Geriatric Dogs Using Thoracic Radiographs

Jiwon Chung
Major in Veterinary Clinical Sciences
Department of Veterinary Medicine
The Graduate School
Seoul National University

Background: Age is a vital information that impacts all facets of veterinary medicine and accurate estimation of age is needed in fields such as emergency and shelter medicine. A few studies involving age estimation based on deep learning of medical images exists in human medicine, however, in dogs automatic age prediction with machine learning has been based solely on non-medical portrait images.

Purpose: This study proposes a deep learning solution for age estimation in

geriatric dogs from thoracic radiographs and sets the stage for future deep learning research involving other medical imaging modalities and for development of age-related biomarkers.

Method: A large dataset of canine thoracic radiographs was utilized to train and test the convolutional neural network's performance in the estimation of age based on performance metric mean absolute error.

Results: The network was able to extract age-related information from thoracic radiographs of geriatric dogs to estimate age with moderate correlation to the ground truth and through the analysis of activation maps, the vertebra was identified as the main region containing age-related information consistent with previously known sites of degenerative changes.

Conclusion: The convolutional neural network's feasibility in geriatric age estimation using canine thoracic radiographs was confirmed by the results of this study and allowed visualization of areas that may contain age-related information deciphered by artificial intelligence.

Keywords: Artificial intelligence, Deep learning, Age, Thoracic radiograph, Dog

Student Number: 2021-23419

Table of Contents

Introduction 1

Materials and methods 5

1. **Dataset** 5
2. **Hardware specification** 8
3. **Training, validation, and testing** 9
4. **Statistical correlation analysis** 11
5. **Activation map quantification** 13

Results 14

Discussion 30

Conclusion 32

References 33

국문초록 43

Introduction

Age is a crucial factor that is considered in all areas of veterinary care that enables practitioners to provide tailored medical care for both young and geriatric patients. For instance, the functional reserves of the major organ systems gradually and irreversibly decline with age and age must be considered prior to anesthesia since patients react differently to stresses and anesthetics (Carpenter et al., 2005). Life expectancy also differs according to age, and age is a major deciding factor for adoption in shelters (DeLeeuw et al., 2010].

When an animal visits a veterinary clinic, a veterinarian is frequently requested to assess the age of the animal which is crucial for forensic veterinary work, for dogs without identification tags, for stray dogs when buying and selling pets, and for identifying specific illnesses (Mohandespour et al., 2021). The primary method for determining a dog's age is based on growth plate analysis in growing young dogs and dental examination in others (Ei. et al., 2018; Lewis et al., 2019; A. Barton et al., 1939), although different age estimation such as an ocular age estimation method with emphasis on ocular reflections has also been reported (Tobias et al., 2000).

In humans, age prediction using the appearance of anatomical components in medical images has been extensively explored for many years. The human forensics researchers have created techniques for estimating age based on developmental changes in teeth (Grover et al., 2012 Kanchan-Talreja et al., 2012),

several bones including the ilium and clavicle (Calce et al., 2011, Wade et al., 2011), and cranial sutures (Bassed et al., 2010, Akhlaghi et al., 2010). Furthermore, owing to recent developments in machine learning, artificial intelligence (AI) is used to forecast a patient's chronological age by examining a large number of radiographs of the dentition, carpus, thorax and others (Milosevic et al., 2022, Kim et al., 2017, Karagyris et al., 2019). The human studies proposed the feasibility of AI machine learned age estimation for cardiovascular aging, and enhanced predictability for cardiovascular mortality than the actual chronological age (Ieki et al., 2021, Raghu et al., 2021).

Deep learning is a relatively new and rapidly expanding area of machine learning that utilizes multi-layered deep neural networks and extensive data to formulate abstraction (LeCun et al., 2015). Deep learning offers numerous potential benefits for the healthcare industry, including greater performance, unsupervised learning, and the ability to handle huge, complicated datasets (Miotto et al., 2017). Convolutional neural networks (CNNs), a kind of deep learning, typically have several convolution layers. The bottom layer can learn the local characteristics of the picture and the top layer can automatically combine these local features and learn the overall features. Therefore, convolutional neural networks are considered ideal for processing medical picture data (Yang et al., 2021). In human medicine, deep learning achieves expert level in tasks such as retinopathy detection, determining bone age, identifying skin cancer, and others (Gulshan et al., 2016,

Esteva et al., 2017, Iglovikov et al., 2018).

Meanwhile, only inputs and outputs are known and therefore, deep learning is considered as a "black box" and it is challenging for researchers to comprehend the precise internal workings of the CNN. A method to help understand the process of AI network's abstraction would therefore be helpful and Gradient-weighted class activation mapping (Grad-CAM), is a visual explanation technique that highlights key areas in the image that are crucial for idea prediction by using the gradients of any target concept flowing into the final convolutional layer (Selvaraju et al., 2020]. Grad-CAM can create a rough localization map that highlights the key areas in the image for concept prediction thereby allowing a partial interpretation of the machine's reasoning process, and therefore has been employed in many studies involving the use of convolutional neural networks including this study (Selvaraju et al., 2020).

In veterinary medicine, AI is just beginning and only a few studies assessed CNN identifying aberrant pulmonary findings in cats, identifying cardiomegaly in dogs, and figuring out the percentage of feline reticulocytes (Dumortier et al., 2022, Burti et al., 2020, Vinicki et al., 2018).

Meanwhile, the estimation of dog age using AI was attempted using an automated method by assessing a database of portrait photographs (Zamansky et al., 2019). The study revealed that the automated method could estimate the age on average within three years of the actual age in 285 dogs. Although age-related changes such

as spinal degenerative changes, costochondral calcification, increased broncho-interstitial patterns are commonly observed on thoracic radiographs in dogs, to authors' knowledge, there is no study to assess the predictability of AI deep learning for the age based on the medical images. Therefore, in this study, a deep learning solution incorporating convolutional neural networks was applied to estimate age of older dogs based on thoracic radiographs, one of the most commonly used imaging in veterinary practice. We hypothesized that AI can predict the age in geriatric dogs with statistically significant accuracy through machine learning of large thoracic radiograph datasets. The purpose of this study was to evaluate the feasibility of AI deep learning for geriatric age estimation and present the areas that drove the network to decide.

Materials and methods

1. Data selection and workflow

In this retrospective, cross sectional design study, the CNN model training and validation, and internal and external test of trained CNN model using thoracic radiograph data obtained from dogs (**Figure 1**). At first, the CNN model training and validation were performed and the model that displayed the best performance was selected for the internal and external testing. The internal and external testing were analyzed with mean absolute error (MAE) value performance metric and statistical correlation analysis. In addition, the activation maps were generated using external testing data to show the region where the algorithm was most interested.

For CNN model training, model validation, and internal testing, thoracic radiograph data were obtained from publicly available dataset from AI-HUB (www.aihub.or.kr). AI-HUB is an integrated platform organized by the National Information Society Agency in the South Korea to provide AI infrastructure required for the development of AI-based technology, products, and services. At first, 156,942 thoracic radiographs were extracted from the AI-HUB dataset. Then, final radiographs for training were selected by excluding the inappropriate data as follows; (1) The duplicated radiographs were removed with duplicate removal software VisiPics V1.3 (http://www.visipics.info/index.php?title=Main_Page). (2) The inappropriately classified data such as incorrect views, abdominal

radiographs, skull radiographs were selected manually by four veterinarians (Y. C., S. L., S. P., S. R.) with 1 year of radiology experience under the supervision of a veterinarian (J. C.) with 2 years of radiology experience. Finally, 21,609 lateral and 14,100 VD radiographs were subjected to the algorithm's training. The total data were split to be used 80% for training, 10% for verification, and 10% for test datasets, respectively. Then, the ratio was adjusted to some extent to ensure a similar age distribution in each dataset. Finally, in each view, 74.83%, 12.58%, and 12.58% of the total VD views, and 77.71%, 11.14%, and 11.14% of the total lateral views were used for the training, validation, and test datasets, respectively. All partitions have same institution and stratified data split provided similar age distribution in each dataset.

The external testing of the model was performed using the lateral and VD views of thoracic radiographs taken at Seoul National University Veterinary Medical Teaching Hospital (SNU VMTH) from 2011-2022. Information such as birth data, image acquisition date, and radiographic position were extracted and chronological age was calculated from birth date and radiograph acquisition date. If either birth or radiograph date was missing, the radiograph was excluded. Finally, the age and radiographic position was labeled. Ground truth was defined as rounded chronological age. The year age was calculated by dividing the day age by 365. The day age was calculated by subtracting the date of birth from the date of radiograph acquisition.

CNN model training, model validation, testing, and performance evaluation were performed using each lateral and VD dataset separately by a veterinarian (H.L.) with 2 years of experience in clinical pathology and 6 years of experience in the

field of computer science using Intel Xeon CPU 2.20GHz, 51.00GB RAM, 167GB solid-state drive, NVIDIA Tesla V100 SXM2 GPU, and Ubuntu 18.04.6 LTS operating system. The activation map analyzed by a veterinarian (J.C.) with 2 years of radiology experience.

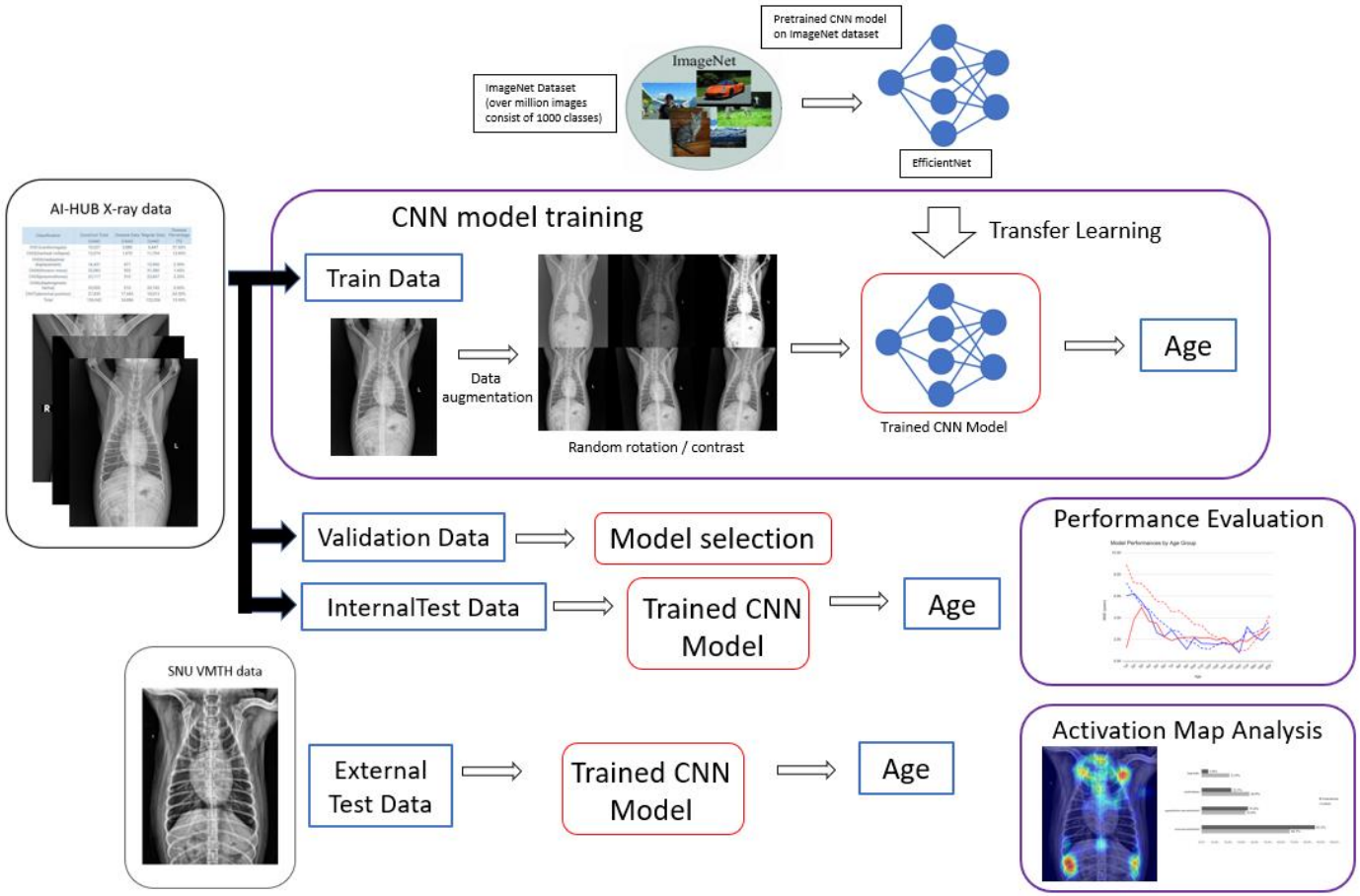

Figure 1. Data usage and overall study framework.

2. Training, validation, and testing

A convolutional neural network-based architecture, EfficientNet (Google AI), was used as deep neural network architecture for canine age estimation models because it showed achievements in various medical AI problems (Marques et al., 2022). EfficientNet, an ImageNet pretrained CNN, is used in transfer learning. EfficientNet was trained using the selected AI-HUB data in the fine-tuning stage through data augmentation which includes various transforms such as random rotation and random contrast. Input image size was 3-channel image of 1024 by 1024 pixels. EfficientNetB0 and EfficientNetB1 were adopted as encoder layer. The encoded features were subjected to decoder layer which consists of two hidden layers which have 512 nodes. Output node has 20 nodes. The output was inferenced as logit because target range of age prediction was from 0 to 20. OpenCV, an image processing library, and PyTorch, an open-source deep learning framework for python was utilized. The model parameter was initialized as ImageNet pretrained for EfficientNet in transfer learning method. Random rotation ranged from -5 degree to +5 degree and random contrast, and random translation were used for image augmentation. In training, preprocessing consisted of contrast limited adaptive histogram equalization, and normalization using mean and standard deviation of pixel intensities.

During training, the preprocessed radiographs and calculated age were fed in each step into EfficientNet. Loss calculation between ground truth and predicted ages, which is required for back-propagation to train numerous parameters of

EfficientNet, was selected as mean variance loss which is widely utilized and showed success in age estimation of humans. Varying size models of the EfficientNet are selected and initial values were adjusted according to the learning rate and augmentation degree.

Multiple models were simultaneously trained and validated after each epoch to find the model that excelled in comparative performance. Every training experiment was designed as training of 100 epochs. An epoch is defined as one-cycle of each training dataset. At each end of the epoch, the EfficientNet was evaluated by a validation dataset using MAE metric. If there was no improvement in MAE in a row, training was prematurely terminated. The best performing model was selected and was tested with both internal and external datasets.

After the end of training, the EfficientNet was evaluated by internal test datasets comprised of 10% of the AI HUB data; 2,719 lateral and 1,571 VD radiographs for internal testing. The external testing data acquired from SNU VMTH was used to gauge the performance of the trained networks on data of different source from that of the training data and 1,015 lateral and 1,023 VD radiographs were used for the external testing.

3. Statistical correlation analysis

Data analyses were performed using Microsoft Office Professional Plus 2019 by (H. L.). Two metrics are used to gauge performance: the mean absolute error and the coefficient of determination using Pearson's coefficient (r). The average discrepancy (in years) between estimated and actual ages is known as the mean absolute error which is an adaptation of standard performance measures in human age estimation (Geng et al., 2007). It is given by the formula [19]:

$$MAE = \frac{\sum_{i=1}^{n} abs\left(y_i - \lambda(x_i)\right)}{n}$$

where y_i is the true target value for test instance x_i , $\lambda(x_i)$ is the predicted target value for test instance x_i , and *n* is the number of test instances.

The coefficient of determination (R^2) was used to measure how well a statistical model predicted an outcome and was given by the following equation:

$$Coefficient\ of\ determination\left(R^2\right) = (r)^2$$
$$where\ r = Pearson\ correlation\ coefficient$$

The outcome was represented by the model's dependent variable. The value was calculated using the most common method for determining a linear connection is the Pearson correlation coefficient (r). The degree and direction of the link between two variables is expressed as a number between -1 and 1. The lowest possible value of R^2 is 0 and the highest possible value is 1. Put simply, the better a model is at making predictions, the closer its R^2 will be to 1.

Geriatric patients are those who have lived 75-80% of their predicted lifespan, in

general (Dodman et al., 1984). Defining threshold age for geriatric patient is equivocal since different dog breeds have different life expectancies. A study of beagles' cognitive capacities classified them as old from 8 years of age, and the old group made mistakes that suggested diminished learning (Tapp et al.,2003). For the purpose of this study, we considered patients above 8 years of age as geriatric.

4. Activation Map Quantification

After gauging performance, gradient-weighted class activation mapping (Grad-CAM) was utilized to create activation maps in order to help understand which regions of the thoracic radiographs drive the network for final estimation. The activation maps enable us to see the regions in which the algorithm is most interested.

A total of 400 activation maps that comply most closely to standardized radiographic positioning were selected by a veterinarian with 2 years of radiology experience (J.C.). The activation regions were classified into 4 regions: lung fields, mediastinum, appendicular musculoskeletal system, and axial musculoskeletal system. The activated regions were tallied by four veterinarians (Y. C., S. L., S. P., S. R.) with 1 year of radiology experience under the supervision of a veterinarian (J.C.) with 2 years of radiology experience. The percentage of regional activations were calculated separately for lateral and VD activation maps for comparison.

Results

Age estimation results

The coefficient of determination was 0.60 and Pearson's correlation coefficient R was 0.78, with p-value below 0.00001. The estimated ages showed moderate correlation with the actual ages (**Figure 2**).

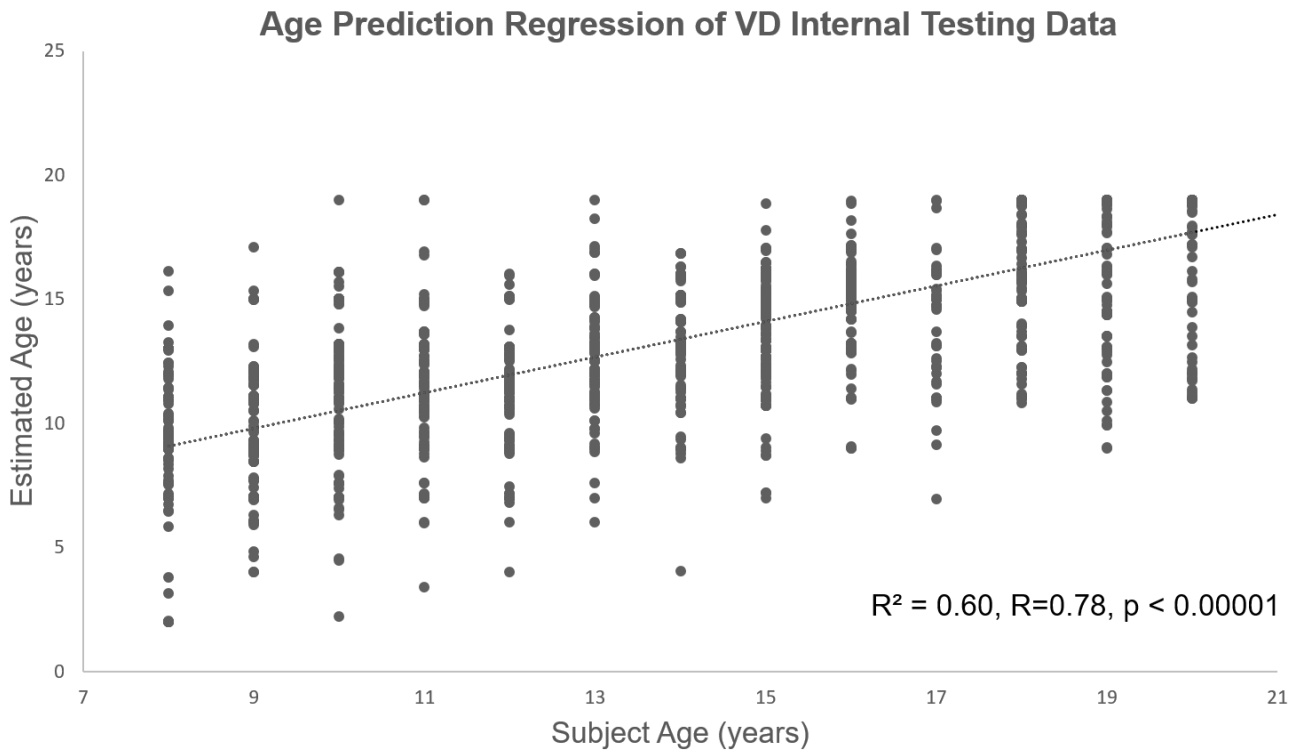

Figure 2 Regression analysis. The estimations given by the algorithm for VD testing data displays moderate correlation to the actual ages.

The algorithm's performance was also measured via MAE values, in which lower values signify better performance. The MAE values for all dogs in the internal test data were below 2 years for VD radiograph-based algorithm and slightly above 2 years for lateral radiograph-based algorithm. In the dog data older than 8 years, the model showed considerably better performance, while it displayed subpar performance in dealing with dog data under the age of six.

Table 1. Mean absolute error of the algorithm's performance for estimating age in the internal test data

Age (Range)	Estimated age on lateral radiographs	Estimated age on VD radiographs
All	2.31	1.83
> 8 years	2.12	1.62
< 6 years	3.44	4.01

Table 1 presents the mean absolute error of the internal test data. The neural network's performance on the external test data for both the VD and lateral were below than that of the internal test data (**Table 2**). The network performed well at a level comparable to that of the internal test data for patient data older than 8 years old. Performance was significantly below average when estimating patients under the age of six. In comparison to the performance on the internal test data, the performance on the external test dataset was comparably below average. **Table 2** presents the mean absolute error of the external test data.

Table 2. Mean absolute error of the algorithm's performance for estimating age in the external test data

Age (Range)	Estimated age on lateral radiographs	Estimated age on VD radiographs
All	3.98	2.94
> 8 years	2.36	1.81
< 6 years	6.75	5.09

Interpretation of the neural network by Grad-CAM analysis

The algorithm's activation maps for lateral and VD thoracic radiographs with age correctly predicted within ± 2 years are presented. In the majority of the activation maps, increased level of attention was evident in the cervical vertebra region regardless of the position **(Figure 3).** Some maps displayed multiple activation regions while others focused primarily on the vertebra **(Figure 4).**

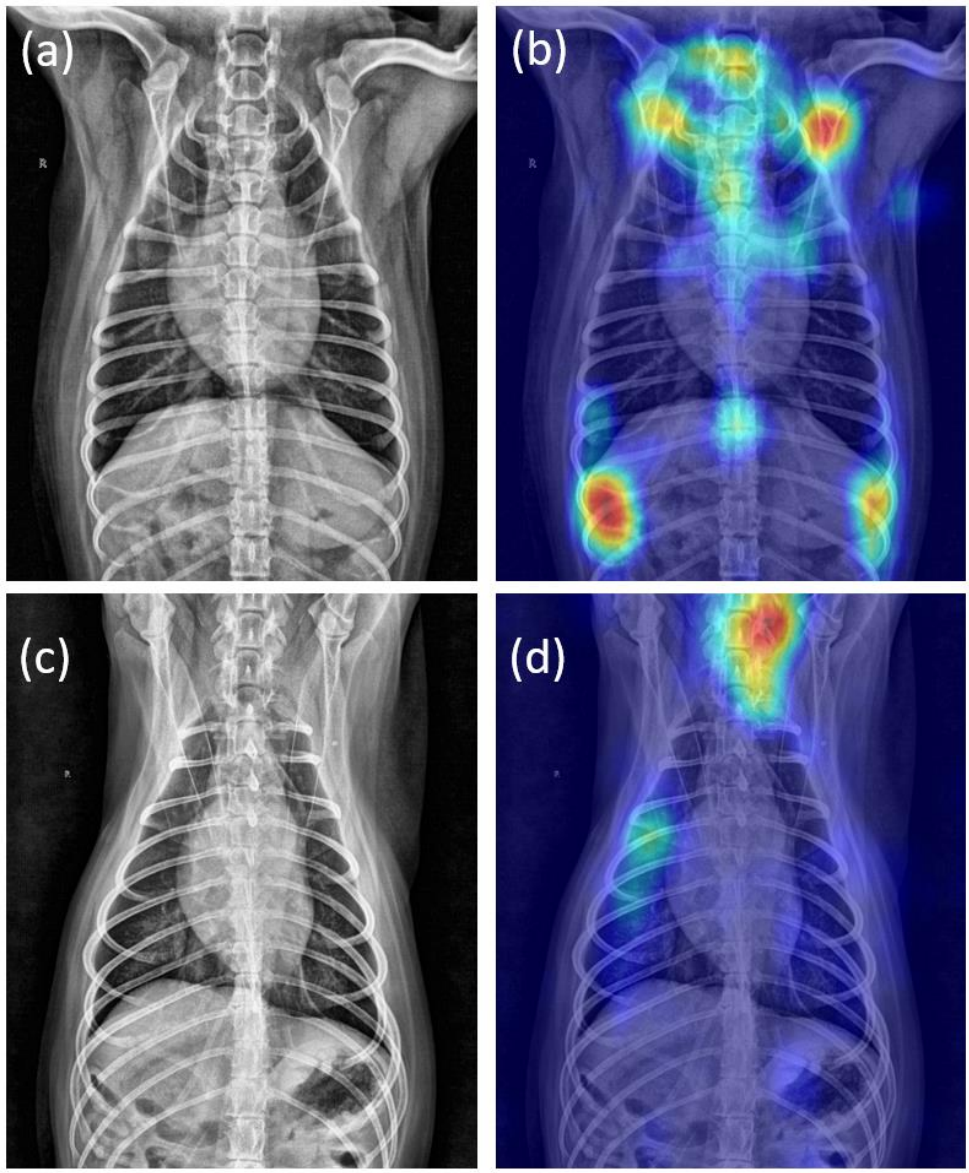

Figure 3. VD activation maps. (a) Radiograph of a 10 year old, castrated male patient with (b) corresponding activation map for age estimation. Activation regions are evident in the cervical-thoracic vertebrae, shoulder joints, ribs, parts of lungs, and mediastinum. (c) Radiograph of an 8 year old, male castrated patient with (d) corresponding activation map for age estimation. Activation regions are evident in the cervical vertebrae and in the right lung field.

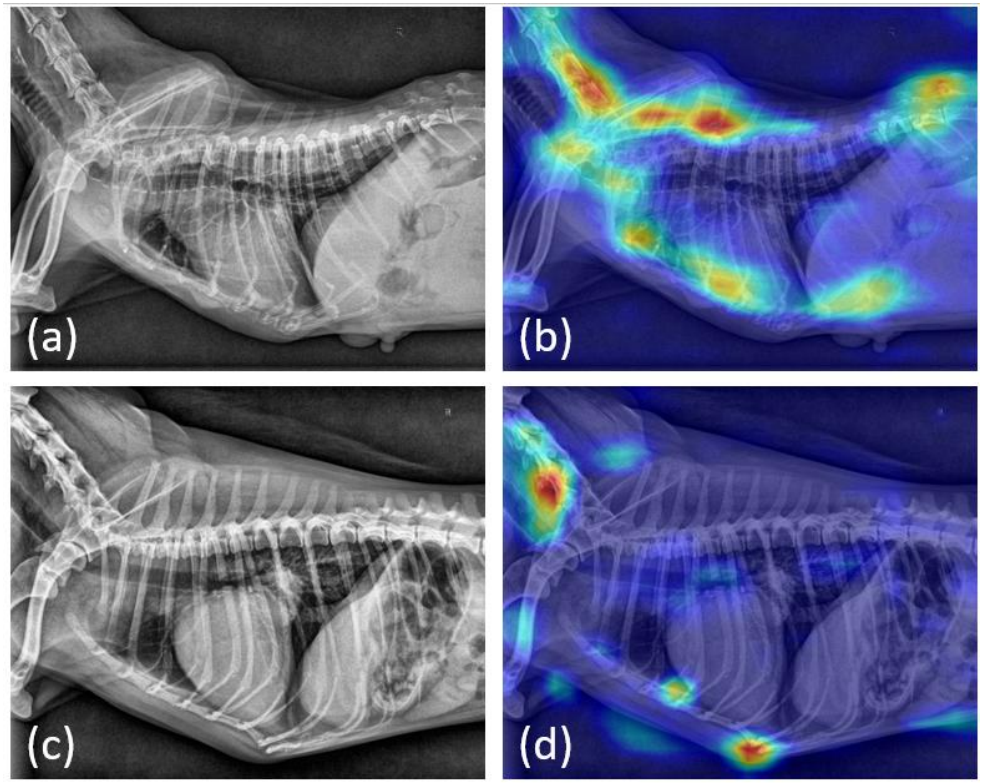

Figure 4. Lateral activation maps. (a) Radiograph of a 17 year old, female patient with (b) corresponding activation map for age estimation. Activation regions are evident in the cervical-thoracic-lumbar vertebrae, sternum, xiphoid, shoulder joints, and mediastinum. (c) Radiograph of a 7 year old, castrated male patient with (d) corresponding activation map for age estimation. Activation regions are evident in the cervical vertebrae, sternum, xiphoid, scapula, and forelimb.

The activation maps can be deciphered by the color spectrum with red signifying the areas of highest levels of attention followed by yellow, green, and blue in an order of decreasing attention levels. On the activation map, the axial skeletal system including the cervical spine, ribs, and sternum, appeared to be the main region of interest to the algorithm for extracting age-related information. Meanwhile, the mediastinum and the lungs were given less consideration.

Compared to the activation map of the lateral radiographs, the analysis of the VD activation maps differed in that while the network displayed principal interest in the vertebra of the axial skeletal system, it also considered regions such as the appendicular skeletal system, mediastinum, and lungs to a larger degree.

Activation maps differed in that some activation maps showed activations across several regions **(Figure 5).** Out of the 400 selected activation maps, most of the algorithm's attention centered around the axial musculoskeletal system for both VD and lateral radiographs.

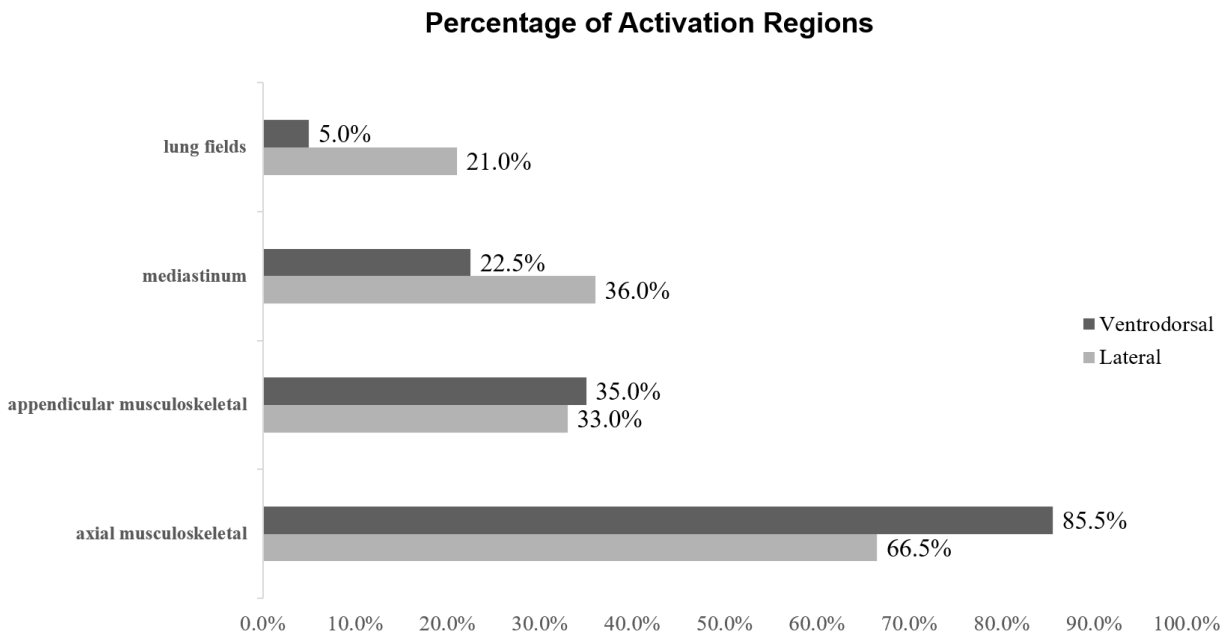

Figure 5. Percentage graph of the counted activation regions in 200 ventrodorsal and 200 lateral activation maps.

The neural network appeared to interpret lateral radiographs with greater heterogeneity than it did VD thoracic radiographs, meaning that it was more interested in analyzing also the mediastinal and lung regions in the latter.

While the neural network's performance differs across different radiographic view positions, it also varied across different age groups (**Figure 6**).

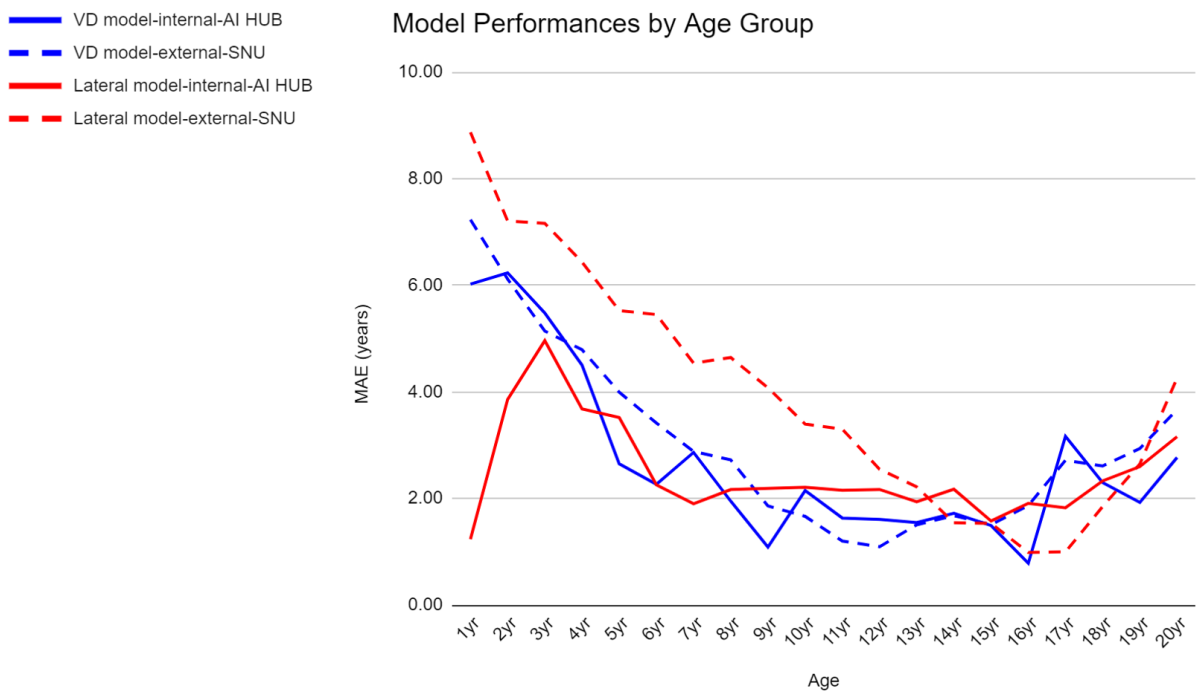

Figure 6. Performance measurements of the algorithms for internal [AI HUB] and external [SNU] test datasets by age group. Error tends to decrease in the groups with increasing age.

Discussion

This study investigated the feasibility of AI deep learning for estimating geriatric canine age using thoracic radiographs. The convolutional neural network could predict ages in dogs with moderate correlation to the ground truth. In particular, the vertebra was the primary area of activation, according to the activation maps. Other regions that the network considered include the mediastinum, lung fields, and appendicular musculoskeletal system. In estimating the ages of geriatric patients older than 8 years of age, the neural network's performance was around MAE 2 years. The algorithms showed variations in estimating performance and in areas of activation, depending on whether the algorithms were fine-tuned using the lateral or VD radiographs.

In this study, the estimated ages and the actual ages displayed moderate correlation. The network's performance in MAE by age group shows that the error values tend to decrease with increasing age with age groups 13-15 years producing estimates close to within 2 years of the actual age on average. It was interesting to note that the algorithm that used VD radiographs as its foundation performed better than when lateral radiographs were used. A study in human medicine revealed that algorithm trained with the anterior-posterior radiographic view position trailed in performance in comparison to that of the posterior-anterior learned algorithm for unknown reasons (Karagyris et al., 2019).

The analysis of the activation mapping showed that age-related information used by the artificial intelligence resides primarily in the vertebra of the axial musculoskeletal system with degenerative changes in the bones most likely being the deciding factor for the network. In the canine vertebra, degenerative changes such as marked discal degeneration, disc space narrowing and collapse, calcification of the discal tissue, and sclerosis of the bony endplates have been described (Resnick et al., 1985, Morgan et al., 1988). The neural network could autonomously recognize these features of vertebral degeneration and make statistical association to the increasing age. Likewise, the shoulder joint degenerative changes such as subchondral bone sclerosis, formation of enthesophytes and periarticular osteophytes, and joint contour remodeling would be learned by the neural network to be associated with older age in older canine thoracic radiographs (Morgan et al., 1987).

The secondary source of age-related information appeared to be within the mediastinum, lung fields, and shoulder regions. In older dogs, radiographic changes of the lungs can be seen such as pleural thickening, increased linear markings, nodular densities, and increased tracheal and bronchial wall densities (Rhodes et al., 1966). In human medicine, senile changes in the mediastinum can be observed radiographically, such as tortuous coronary arteries and vascular and valvular calcifications, aortic modifications, and tracheal appearance changes as well as in the thoracic wall in the form of reduced thickness (Redheuil et al., 2011, Reif et al., 1966). Given that the network performance was superior and that more mediastinum and lung region activations were seen utilizing the VD radiographs, it is feasible to assume that additional age-related information is available in regions

other than the vertebra which was used by the algorithm to enhance performance.

This study was limited in that the breed information was not available. We assumed that the majority of the thoracic radiographs obtained were from small dog breeds prominent in S. Korea and estimating the age of the mid- to large-breed dogs using the neural network used in this study should be investigated further.

In human medicine artificial intelligence age estimation is explored in other medical medical images such as the hand radiographs and brain MRI's (S.Wang et al., 2018, J.Wang et al., 2019). Likewise, similar applications may be feasible in veterinary medicine such as to accurately estimate the age of growing young dogs and to use as a biomarker for aging-related brain diseases. Also, since exact time of death information is available in people, age estimation from chest radiographs showed potential to be a biomarker to predict cardiovascular disease related mortality (Ieki et al., 2021). Maintaining accurate time of death as well as date of birth information in veterinary medicine and its incorporation in future artificial intelligence research may lead to development of more feasible and accessible biomarkers to predict longevity.

Conclusion

This study assessed the feasibility of training a convolutional neural network to estimate age in geriatric canine patients based on thoracic radiographs. The artificial intelligence was able to produce age estimates with moderate correlation to the true ages. In addition, activation maps identified regions such as the vertebra, the mediastinum, and the lungs to contain age-related degenerative information extracted by the artificial intelligence. Since degenerative changes are not normally apparent in young patients, other types of radiographs such as carpal radiographs to assess growth plates can be considered for similar further studies. As this study is the first work in automatic age estimation in veterinary medicine using convolutional neural networks and medical images, we want to set the stage for future research directions. Future research involving various other imaging modalities combined with deep learning and additional patient metadata including time of death, date of diagnoses, and accurate date of birth may lead to development of age-related biomarkers to predict longevity, monitor health status, and develop appropriate management plans for veterinary patients.

References

Akhlaghi, M., Taghaddosinejad, F., Sheikhazadi, A., Valizadeh, B., & Rezazadeh Shojaei, S. M. (2010). Age-at-death estimation based on the macroscopic examination of Spheno-occipital sutures. Journal of Forensic and Legal Medicine, 17(6), 304–308. https://doi.org/10.1016/j.jflm.2010.04.009

Ei, O., Yakubu A., Kene R., & Shehu S. (2018). Radiographic Evaluation of the Appearance and Closure Time of Growth Plates of Radius and Ulna Bones in Nigerian Indigenous Dogs. Journal of Veterinary and Animal Research, 1(1), 1–6. https://doi.org/10.18875/2639-7315.1.101

Bassed, R. B., Briggs, C., & Drummer, O. H. (2010). Analysis of time of closure of the spheno-occipital synchondrosis using computed tomography. Forensic Science International, 200(1–3), 161–164. https://doi.org/10.1016/j.forsciint.2010.04.009

Bonomo, L., Larici, A. R., Maggi, F., Schiavon, F., & Berletti, R. (2008). Aging and the Respiratory System. Radiologic Clinics of North America, 46(4), 685–702. https://doi.org/10.1016/j.rcl.2008.04.012

Burti, S., Longhin Osti, V., Zotti, A., & Banzato, T. (2020). Use of deep learning to detect cardiomegaly on thoracic radiographs in dogs. Veterinary Journal, 262, 105505. https://doi.org/10.1016/j.tvjl.2020.105505

Calce, S. E., & Rogers, T. L. (2011). Evaluation of age estimation technique: Testing traits of the acetabulum to estimate age at death in adult males. Journal of Forensic Sciences, 56(2), 302–311. https://doi.org/10.1111/j.1556-4029.2011.01700.x

Carpenter, R. E., Pettifer, G. R., & Tranquilli, W. J. (2005). Anesthesia for geriatric patients. Veterinary Clinics of North America - Small Animal Practice, 35(3), 571–580. https://doi.org/10.1016/j.cvsm.2004.12.007

Chandrakanth, H. V., Kanchan, T., Krishan, K., Arun, M., & Kumar, G. N. P. (2012). Estimation of age from human sternum: An autopsy study on a sample from South India. International Journal of Legal Medicine, 126(6), 863–868. https://doi.org/10.1007/s00414-012-0752-0

Deleeuw, J. L. (2010). Animal shelter dogs: factors predicting adoption versus euthanasia a dissertation by Jamie L. DeLeeuw Master of Arts, Wichita State University , 2008 Bachelor of Science , Grand Valley State University , 2004 Submitted to the Department of Psychology a. December.

Dumortier, L., Guépin, F., Laure, M., & Muller, D. (2022). Deep learning in veterinary medicine , an approach based on CNN to detect pulmonary abnormalities from lateral thoracic radiographs in cats. Scientific Reports, 1–12. https://doi.org/10.1038/s41598-022-14993-2

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118. https://doi.org/10.1038/nature21056

Gilsanz, V., & Ratib, O. (2012). Hand Bone Age Bone Development. In Hand Bone Age. http://link.springer.com/10.1007/978-3-642-23762-1

Grover, S., Marya, C. M., Avinash, J., & Pruthi, N. (2012). Estimation of dental age and its comparison with chronological age: Accuracy of two radiographic methods. Medicine, Science and the Law, 52(1), 32–35. https://doi.org/10.1258/msl.2011.011021

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA - Journal of the American Medical Association, 316(22), 2402–2410. https://doi.org/10.1001/jama.2016.17216

Ieki, H., Ito, K., Saji, M., Kawakami, R., Nagatomo, Y., Koyama, S., Matsunaga, H., Miyazawa, K., Ozaki, K., Onouchi, Y., Katsushika, S., Matsuoka, R., Shinohara, H., Yamaguchi, T., Kodera, S., Higashikuni, Y., Fujiu, K., Akazawa, H., Isobe, M., … Komuro, I. (2021). Deep learning-based chest X-ray age serves as a novel biomarker for cardiovascular aging. BioRxiv, 2021.03.24.436773. http://biorxiv.org/content/early/2021/03/25/2021.03.24.436773.abstract

Iglovikov, V. I., Rakhlin, A., Kalinin, A. A., & Shvets, A. A. (2018). Paediatric bone age assessment using deep convolutional neural networks. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11045 LNCS, 300–308. https://doi.org/10.1007/978-3-030-00889-5_34

Kanchan-Talreja, P., Acharya, A. B., & Naikmasur, V. G. (2012). An assessment of the versatility of Kvaal's method of adult dental age estimation in Indians. Archives of Oral Biology, 57(3), 277–284. https://doi.org/10.1016/j.archoralbio.2011.08.020

Karargyris, A., Kashyap, S., Wu, J. T., Sharma, A., Moradi, M., & Syeda-Mahmood, T. (2019). Age prediction using a large chest x-ray dataset. 66. https://doi.org/10.1117/12.2512922

Kim, J. R., Lee, Y. S., & Yu, J. (2015). Assessment of bone age in prepubertal healthy korean children: Comparison among the korean standard bone age chart,

greulich-pyle method, and tanner-whitehouse method. Korean Journal of Radiology, 16(1), 201–205. https://doi.org/10.3348/kjr.2015.16.1.201

Kim, J. R., Shim, W. H., Yoon, H. M., Hong, S. H., Lee, J. S., Cho, Y. A., & Kim, S. (2017). Computerized bone age estimation using deep learning-based program: Evaluation of the accuracy and efficiency. American Journal of Roentgenology, 209(6), 1374–1380. https://doi.org/10.2214/AJR.17.18224

Kim, W. H., Min, K. D., Cho, S. Il, & Cho, S. (2020). The Relationship Between Dog-Related Factors and Owners' Attitudes Toward Pets: An Exploratory Cross-Sectional Study in Korea. Frontiers in Veterinary Science, 7(August), 1–11. https://doi.org/10.3389/fvets.2020.00493

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Lewis, G. (2019). Musculoskeletal Development of the Puppy: Birth to Twelve Months. Animal Therapy Magazine, 15, 41–44. https://www.researchgate.net/publication/333118590_Musculoskeletal_Development_of_the_Puppy_Birth_to_Twelve_Months

Liang, Y., Li, S., Yan, C., Li, M., & Jiang, C. (2021). Explaining the black-box model: A survey of local interpretation methods for deep neural networks. Neurocomputing, 419, 168–182. https://doi.org/10.1016/j.neucom.2020.08.011

Marques, G., Ferreras, A., & de la Torre-Diez, I. (2022). An ensemble-based approach for automated medical diagnosis of malaria using EfficientNet. Multimedia Tools and Applications, 81(19), 28061–28078. https://doi.org/10.1007/s11042-022-12624-6

Milošević, D., Vodanović, M., Galić, I., & Subašić, M. (2022). Automated estimation of chronological age from panoramic dental X-ray images using deep learning. Expert Systems with Applications, 189(October 2021), 116038. https://doi.org/10.1016/j.eswa.2021.116038

Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2017). Deep learning for healthcare: Review, opportunities and challenges. Briefings in Bioinformatics, 19(6), 1236–1246. https://doi.org/10.1093/bib/bbx044

Mohandespour, F., & Mosallanejad, B. (2021). Evaluation of age-related changes in dentition of terrier dogs Evaluation of age-related changes in dentition of terrier dogs. 11196205(January). https://doi.org/10.22055/IVJ.2020.244093.2291

Morgan, J. P., & Miyabayashi, T. (1988). Degenerative changes in the vertebral column of the dog. A review of radiographic findings. Veterinary Radiology, 29(No. 2), 72–77. https://doi.org/10.1201/9781498713283-84

Morgan, J. P., Pool, R. R., & Miyabayashi, T. (1987). Primary degenerative joint disease of the shoulder in a colony of beagles. Journal of the American Veterinary

Medical Association, 190(5), 531—540.

http://europepmc.org/abstract/MED/3558090

Raghu, V. K., Weiss, J., Hoffmann, U., Aerts, H. J. W. L., & Lu, M. T. (2021). Deep Learning to Estimate Biological Age From Chest Radiographs. JACC: Cardiovascular Imaging, 14(11), 2226–2236. https://doi.org/10.1016/j.jcmg.2021.01.008

Redheuil, A., Yu, W. C., Mousseaux, E., Harouni, A. A., Kachenoura, N., Wu, C. O., Bluemke, D., & Lima, J. A. C. (2011). Age-related changes in aortic arch geometry: Relationship with proximal aortic function and left ventricular mass and remodeling. Journal of the American College of Cardiology, 58(12), 1262–1270. https://doi.org/10.1016/j.jacc.2011.06.012

Reif, J. S., & Rhodes, W. H. (1966). The Lungs of Aged Dogs: A Radiographic-Morphplogic Correlation. American Veterinary Radiology Society National Meeting, 2, 5–11.

Resnick, D. (1985). Degenerative diseases of the vertebral column. Radiology, 156(1), 3–14. https://doi.org/10.1148/radiology.156.1.3923556

Sammut, C., Webb, G.I. (2011). Mean Absolute Error, (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-30164-8_525

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2020). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. International Journal of Computer Vision, 128(2), 336–359. https://doi.org/10.1007/s11263-019-01228-7

Schmeling, A., Schulz, R., Reisinger, W., Mühler, M., Wernecke, K. D., & Geserick, G. (2004). Studies on the time frame for ossification of the medial clavicular epiphyseal cartilage in conventional radiography. International Journal of Legal Medicine, 118(1), 5–8. https://doi.org/10.1007/s00414-003-0404-5

Tapp, P. D., Siwak, C. T., Estrada, J., Head, E., Muggenburg, B. A., Cotman, C. W., & Milgram, N. W. (2003). Size and reversal learning in the beagle dog as a measure of executive function and inhibitory control in aging. Learning and Memory, 10(1), 64–73. https://doi.org/10.1101/lm.54403

Tobias, G., Tobias, T. a, & Abood, S. K. (2000). Estimating age in dogs and cats using ocular lens examination. Comp Contin Edu Pract Vet, 22(12), 1085.

Vinicki, K., Ferrari, P., Belic, M., & Turk, R. (2018). Using Convolutional Neural Networks for Determining Reticulocyte Percentage in Cats. http://arxiv.org/abs/1803.04873

Wade, A., Nelson, A., Garvin, G., & Holdsworth, D. W. (2011). Preliminary radiological assessment of age-related change in the trabecular structure of the

human os pubis. Journal of Forensic Sciences, 56(2), 312–319. https://doi.org/10.1111/j.1556-4029.2010.01643.x

Wang, J., Knol, M. J., Tiulpin, A., Dubost, F., De Bruijne, M., Vernooij, M. W., Adams, H. H. H., Ikram, M. A., Niessen, W. J., & Roshchupkin, G. V. (2019). Gray matter age prediction as a biomarker for risk of dementia. Proceedings of the National Academy of Sciences of the United States of America, 116(42), 21213–21218. https://doi.org/10.1073/pnas.1902376116

Yang, S., Zhu, F., Ling, X., Liu, Q., & Zhao, P. (2021). Intelligent Health Care: Applications of Deep Learning in Computational Medicine. Frontiers in Genetics, 12(April), 1–21. https://doi.org/10.3389/fgene.2021.607471

Zamansky, A., Sinitca, A. M., Kaplun, D. I., Dutra, L. M. L., & Young, R. J. (2019). Automatic estimation of dog age: The dogage dataset and challenge. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11729 LNCS(September), 421–426. https://doi.org/10.1007/978-3-030-30508-6_34

국문 초록

흉부 방사선 영상을 이용한 딥러닝 기반 노령견 나이 추정에 관한 연구

서울대학교 대학원

수의학과 임상수의학 전공

정 지 원

연령은 수의학의 모든 측면에 영향을 미치는 중요한 정보이며, 이 연구는 개 흉부 방사선 영상을 사용하여 노령견들의 연령 추정을 하는 deep learning solution 을 제안합니다. 이 연구는 나이 추정을 위해 컨볼루션 신경망을 학습시키고 성능을 시험하기 위해 공개적으로 사용 가능한 대규모 방사선 데이터 세트와 동물 의료 기관의 추가 데이터를 사용하여 수행되었습니다. 이 네트워크는 흉부 방사선 영상에서 연령 관련 정보를 추출하여 상당히 정확하게 노령견에서 연령들을 추정할 수 있었습니다. 또한 활성화 맵 분석을 통해 축근골격계가 연령 관련 정보를 포함하는 주요 영역으로 식별되었으며 이 결과는 이전

연구들에서 알려졌던 축근골격계에서 나타나는 퇴행성 변화들을 네트워크가 인식했을 가능성을 제시합니다. 흉부 방사선 촬영은 여전히 호흡기 및 심혈관 질환을 선별하기 위해 수의학에서 가장 자주 사용되는 영상 검사 중 하나이기 때문에 동물보호소 의학, 예방 의학 및 반려 동물 보험과 같은 영역에서 자동 연령 추정의 실제 사용을 기대할 수 있습니다. 본연구는 딥 러닝과 의료 영상을 연령 추정에 활용해본 수의학 분야 최초의 연구로서, 향후 소동물의학 분야에서 컨볼루션 신경망을 활용한 연령 측정 및 연령 관련 바이오마커 개발 연구등의 발판을 마련하고자 합니다.

주요어: 인공지능, 흉부 방사선, 딥 러닝, 나이, 개

학 번: 2021-23419