



Article Bone Metastasis Detection in the Chest and Pelvis from a Whole-Body Bone Scan Using Deep Learning and a Small Dataset

Da-Chuan Cheng ^{1,2,*}, Chia-Chuan Liu ³, Te-Chun Hsieh ^{1,2,4}, Kuo-Yang Yen ^{1,4} and Chia-Hung Kao ^{2,4,5,6,*}

- ¹ Department of Biomedical Imaging and Radiological Science, China Medical University, Taichung 404, Taiwan; D10119@mail.cmuh.org.tw (T.-C.H.); T10540@mail.cmuh.org.tw (K.-Y.Y.)
- ² Center of Augmented Intelligence in Healthcare, China Medical University Hospital, Taichung 404, Taiwan
- ³ Department of Medical Image, Taipei Medical University-Shuang Ho Hospital, New Taipei 235, Taiwan; 20159@s.tmu.edu.tw
- Department of Nuclear Medicine and PET Center, China Medical University Hospital, Taichung 404, Taiwan
- ⁵ Department of Bioinformatics and Medical Engineering, Asia University, Taichung 413, Taiwan
- ⁶ Graduate Institute of Biomedical Sciences and School of Medicine, College of Medicine, China Medical University, Taichung 404, Taiwan
- * Correspondence: dccheng@mail.cmu.edu.tw (D.-C.C.); d10040@mail.cmuh.org.tw (C.-H.K.)

Abstract: The aim of this study was to establish an early diagnostic system for the identification of the bone metastasis of prostate cancer in whole-body bone scan images by using a deep convolutional neural network (D-CNN). The developed system exhibited satisfactory performance for a small dataset containing 205 cases, 100 of which were of bone metastasis. The sensitivity and precision for bone metastasis detection and classification in the chest were 0.82 ± 0.08 and 0.70 ± 0.11 , respectively. The sensitivity and specificity for bone metastasis classification in the pelvis were 0.87 ± 0.12 and 0.81 ± 0.11 , respectively. We propose the use of hard example mining for increasing the sensitivity and precision of the chest D-CNN. The developed system has the potential to provide a prediagnostic report for physicians' final decisions.

Keywords: bone metastasis; deep learning; hard example mining

1. Introduction

According to a report published in 2018 by the National Health Insurance Research Database of Taiwan, prostate cancer (PC) is the seventh highest ranking cause of cancerrelated deaths among Taiwanese men [1]. PC has a high degree of osteotropism [2] because the possibility of metastases is relatively high; however, PC has a slower progression than many other cancers. According to the American Cancer Society, if PC has only spread to the bones and not to other organs, radium-223 can be used to help people live longer [3]. If the cancer has grown outside the prostate, preventing or slowing the spread of the cancer to the bones is a major treatment goal. If the cancer has already reached the bones, controlling or relieving pain and other complications is an important part of treatment. The five-year relative survival rate for individuals with PC that has spread to distant lymph nodes, organs, or the bones is 29% [3]. Patients with only bone metastases can be treated with hormone therapy, chemotherapy, or radiation therapy. Early identification of PC metastases is important because therapy can effectively slow metastasis progression at this stage. One of the current diagnostic media used in clinics for bone metastasis diagnosis is the whole-body bone scan (WBBS), with the vein injection of the Tc-99m MDP tracer. The aim of this research was to develop an automated system for helping physicians to detect bone metastasis in the early stage. In this study, we propose two neural network (NN)-based systems: (1) an D-CNN-based deep learning technique that can identify bone metastases in the pelvis as early as possible, and (2) a faster region-based convolutional NN



Citation: Cheng, D.-C.; Liu, C.-C.; Hsieh, T.-C.; Yen, K.-Y.; Kao, C.-H. Bone Metastasis Detection in the Chest and Pelvis from a Whole-Body Bone Scan Using Deep Learning and a Small Dataset. *Electronics* **2021**, *10*, 1201. https://doi.org/10.3390/ electronics10101201

Academic Editor: Jian Sun

Received: 10 April 2021 Accepted: 14 May 2021 Published: 18 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). (R-CNN) that can identify metastasis spots on the ribs or spinal cord, if any, in the WBBS. Both these systems aim to help a physician in the early detection of small metastasis.

Small metastases can also be identified by measuring the bone scan index (BSI). The BSI was proposed in 1998 [4]. A US patent related to the BSI was issued in 2012 [5]. The related publication of this patent is [6]; however, no description of the measurement technique was provided in the publication. We only know that in [6], the authors extracted 20–30 features of the hotspots and used an NN as a classifier. They used 795 patients as the training group, and the number of hotspots collected was >40,000 for various metastatic cancers (e.g., prostate, breast, and kidney cancer). The system used in [6] suitably detected hotspots in certain areas; however, it could not detect hotspots in the large area of bone metastasis (Figure 3 in [6]). The reason for this result might be that the training data on hotspots were limited. To the best of our knowledge, no researchers have described a technique for bone metastasis detection or identification. In [7], the authors used ResNet50 as a backbone and incorporated the ladder network to form a ladder feature pyramid network (LFPN), which can use unlabeled data for bone metastasis detection. The mean sensitivity and precision of lesion detection were 0.856 and 0.852, respectively. For metastasis classification (four classes) in the chest, the sensitivity and specificity were 0.657 and 0.857, respectively. The aforementioned study provides useful technical details on metastasis detection and classification by using deep learning.

The remainder of this paper is organized as follows. The image resources, difficulties, and models are described in Section 2. The results are presented in Section 3. A discussion of the results is provided in Section 4, and the conclusions are detailed in Section 5.

2. Materials and Methods

2.1. Materials

No patient preparation was required. Patients underwent whole-body planar bone scans with a gamma camera (Millennium MG, Infinia Hawkeye 4, or Discovery NM/CT 670 system; GE Healthcare, Waukesha, WI, USA). Bone scans were acquired 2–6 h after the intravenous administration of 20 mCi of technetium-99m methylene diphosphonate (Tc-99m MDP) by using a low-energy high-resolution or general-purpose collimator with a matrix size of 1024×256 , an acquisition time of 15-20 cm/min, and the photon energy centered on the 140 keV photopeak with a symmetrical 20% energy window. During the waiting time and immediately before the scanning, the patients were encouraged to hydrate and void frequently. The patients were scanned in the supine position and whole-body anterior–posterior images were acquired for interpretation. All the images were interpreted using a dedicated GE Xeleris workstation (GE Medical Systems, Haifa, Israel; version 2.0551).

In this retrospective analysis, 205 WBBS images were collected from China Medical University Hospital between August 2013 and May 2019. This study was approved by the Institutional Review Board of China Medical University and Hospital Research Ethics Committee (CMUH106-REC2-130). All the images were studied by two experienced physicians. Hotspots were categorized into two types: (1) confirmed metastatic (or positive) hotspots, and (2) non-cancerous lesions (including degenerative changes and inflammation), and injury (post-trauma). The hotspot classification was confirmed and agreed upon in consensus by the two experienced nuclear medicine physicians according to the available pathological examination, relevant medical history, characteristic findings on other advanced medical imaging modalities (e.g., computed tomography or magnetic resonance image) and/or serial changes observed in follow-up bone scans. All the 205 patients were PC patients. Of the 205 patients, 110 had bone metastasis (confirmed by physicians), and the remaining 95 did not have bone metastasis. To make the detection of hotspots with a computer algorithm easier, we divided the human body into five parts likely to exhibit bone metastasis: the (1) shoulder, (2) rib, (3) spinal cord, (4) pelvis, and (5) thigh. The number and position of the hotspots are summarized in Table 1.

	Shoulder	Rib	Spinal Cord	Pelvis	Thigh
Confirmed metastasis	315	602	399	311	198
Normal lesion	31	81	39	25	11

Table 1. Number and position of the metastasis hotspots.

The patients were aged between 51 and 92 years, and the average age was 73.9 \pm 8.32 years.

The collected WBBS images were in DICOM format, and all private connections were removed. The spatial resolution of the raw image was 1024×512 pixels. The intensity information of each pixel was saved in 2 bytes.

2.2. Difficulties in Bone Metastasis Detection

A difficulty in bone metastasis detection is the differentiation of normal and metastasis hotspots. As injury and osteoarthritis may also cause hotspots, differentiating normal, abnormal, and metastasis hotspots is difficult. For example, injury hotspots may occur in not only one spot but along some spots in a straight line (on ribs). Osteoarthritis hotspots might be symmetric on both sides (left and right). Human experts use certain knowledge to recognize and differentiate hotspots. Such knowledge is non-trivial for mathematically teaching computers or being embedded into an algorithm in traditional image processing techniques. Finding efficient features for classification or object detection is especially difficult.

Radiomics is a traditional method of extracting hand-crafted features [8,9]. Some parts of [6] were also based on radiomics. The CNN has been used for more than 10 years to extract features [7,10,11]. The artificial NN provides an alternative method to extract features. Many studies have indicated that the CNN can extract efficient features automatically during the training phase from numerous training images. In this study, we used the faster R-CNN [12] and YOLO v3 [13,14] state-of-the-art techniques to identify hotspots. In contrast to the CNN, the R-CNN and YOLO techniques can be used for more than simply the classification of an image object. These techniques can be used to detect many objects of interest in an image. The major differences between these two models is that the R-CNN series is a two-stage model, while the YOLO series is a one-stage model. We pick them as representatives for comparison. In the aforementioned techniques, bounding boxes are used to identify the positions of objects and to classify them instantly (i.e., instance segmentation). This property is suitable in metastasis hotspot detection and identification. Figure 1 displays some obtained WBBS images. Figure 1a exhibits no bone metastasis, and Figure 1b contains some metastasis hotspots.





Figure 1. WBBS images: (a) the patient has no bone metastasis and (b) the patient has some bone metastasis hotspots on the rib and spinal cord. The pelvis has a large-area metastasis.

PC might first invade the pelvis and then other sites. Cancer cells may invade the ribs or spinal cord first. To achieve the goal of early bone metastasis detection, we developed two D-CNNs: (1) the D-CNN for pelvis bone metastasis detection (named as the pelvis NN) and (2) the D-CNN for rib and spinal cord bone metastasis hotspot detection (named as the chest NN). This study was divided into the following stages: (1) image enhancement and normalization, (2) detection of five body parts, (3) use of the pelvis NN, and (4) use of the chest NN. In the pelvis NN, only the presence of bone metastasis (yes or no) was determined; however, in the chest NN, the presence of bone metastasis as well as the position of metastatic hotspots in the ribs and spinal cord (both segmentation and classification) were determined.

2.3. Image Preprocessing and Normalization

The normalization of the image size and intensity is an important step prior to image processing. The acquired WBBS images had large variations in the intensity distribution. These variations may have been caused by many factors, such as the blood supply of bones, skeletal (bone) metabolic status, body weight, drug metabolism rate, and leakage of the radiotracer. Some WBBS images had a suitable intensity distribution; however, some other images had poor quality. To alleviate this problem, which might cause problems in image processing, we propose an automated image normalization strategy. This strategy involves image size and image intensity normalization, which are fully automatic.

2.3.1. Spatial Normalization

A standard WBBS image has two views, namely anterior and posterior views. The body range is detected using projection profiles, and both views are cut and centered into an image with a size of 512×950 pixels without scaling or any other transformation. This normalization process has no exception case that does not meet the condition. The normalized image is named as f(r, c).

2.3.2. Intensity Normalization

The intensity of the WBBS images revealed the absorption of Tc-99m MDP by the gamma camera. Leakage of the radiotracer and from the urine bag (usually near the femur) also caused variations in the image intensity. The best method of solving this problem was to focus on the visibility of the tibia and ignore the remaining body parts. A projection profile was created along the *x*-axis. The head projection was on the left part of the profile, and the leg projection was on the right part. A simple algorithm was created to detect two local peaks from the right to the left on the right-hand part. By using the aforementioned strategy, the tibia region can be correctly detected. A linear enhancement is then applied to the tibia region only (from the knee to the foot) as follows:

$$b^* = b \in int(1, \max(f)),$$
$$g(r, c) = 255 \frac{f(r, c)}{b}, if \frac{|k(r, c) \ge a|}{|k(r, c) > 10|} \ge Th$$

where *int*() converts the number to an integer, k(r,c) is the tibia region, a = 50, |*| denotes the count number satisfying the "*" condition, and Th = 0.085 is a percentage threshold. (*r*,*c*) is the coordinate representing the row and column. The parameter *b* is increased from 1 until the "*if*" condition is satisfied. g(r, c) is the intensity normalized image. The aforementioned enhancement process is performed for image intensity normalization. We use the raw data (DICOM format) of the WBBS and convert the image to the PNG format as the input of the D-CNN after image intensity normalization.

2.3.3. Data Augmentation

A large dataset is crucial for achieving a suitable deep learning performance. However, in this study, we only had a small dataset. Therefore, data augmentation was performed to

improve the model performance. Many methods, such as scaling, shearing, rotating, and mirroring, can be used for data augmentation. The intensity normalization procedure can produce one image. According to this normalized image, we can create seven images with different contrast levels. Let g_{max} denote the maximal number in the image. Between g_{max} and b^* , six zones are separated. The length of each zone is $z = (g_{max} - b^*)/7$. Furthermore, linear transformation is used to produce an additional seven contrast images via letting $b = b^* + z$, $b^* + 2z$, until $b^* + 7z$; the transformation is as follows:

$$g(\mathbf{r}, \mathbf{c}) = \begin{cases} 255 \frac{f(r,c)}{b}, & \text{if } f(r,c) < b\\ 255, & \text{otherwise} \end{cases}$$

Another augmentation method is mirroring. In this method, the anterior and posterior views are simply mirrored to double the data number.

2.4. Detection of Five Body Parts

We modified the faster R-CNN to a light version. As this network was only used to detect the five body parts, we resized the image to 160×200 for the input layer. The output layer comprises the label and its bounding box. We only selected the best bounding box for each class and then performed mirror mapping for its anterior or posterior view. In this study, we used only the chest and pelvis parts as the initial sub-images for next-stage inputs, such as chest NN and pelvis NN. The network structure is displayed in Figure 2. The reason why we still need to detect five parts is as follows. The five parts—the shoulders, ribs, spinal cord, pelvis, and thighs—have a tendency towards having cancerous bone metastases of prostate and breast cancer. Our final goal is to detect all lesions on these five parts, although this study focused only on the chest region (including the shoulders, ribs, and spinal cord) and pelvis.



Figure 2. D-CNN for the detection of five body parts. The proposal layer provides bounding boxes (ROI) for the ROI pooling to feed forward to its following CNN, which functions as a classifier.

2.5. Pelvis NN

We only examined whether the pelvis part had bone metastasis; therefore, the output class had only two categories: yes and no, as shown in Figure 3. We used three CNNs as the backbone and modified them in the final fully connected (fc) layer. The NN used for the detection of the five body parts can identify the pelvis part and combine the anterior and posterior views to form a two-view image as an input image for the pelvis NN. The input image size was fixed as $112 \times 287 \times 1$ pixels. The NN used for the detection of five body parts might output pelvis images of different sizes, and the two-view image is resized to fit the input size of the pelvis NN. In the resizing action, the same scaling is used in the *x*-direction and *y*-direction. The remaining part is padded to zero. The resizing action will change the original resolution, and different patients have different scaling factors because their pelvises are different sizes. However, the CNNs are used to recognize if there is any bone metastasis; therefore, the change in original pixel size does not play an important role in this stage. Ten-fold cross validation was performed to calculate their sensitivities and specificities.



Figure 3. Pelvis NN for metastasis classification. We applied three CNN models for comparisons.

2.6. Chest NN

The goals of the chest NN are to detect the positions of hotspots and to classify the hotspots (normal or metastasis). To achieve these goals, we compared two state-of-the art methods, namely the faster R-CNN and YOLO v3. The input layer of the chest NN had a fixed size of $346 \times 292 \times 3$ pixels. The output layer was of two types: (1) one type comprised three classes and (2) the other type comprised bounding boxes. We designed a light version of the faster R-CNN for users possessing a single Nvidia GTX 1080 Ti graphic card. The network structure is displayed in Figure 4. The applied YOLO v3 was from the original network [13,14] without change.

The input layer had three dimensions. The first and second images were the anterior and posterior views of the chest, respectively. The third image, B(r,c), was a nonlinear combination of the anterior and posterior images by $B(r,c) = R(r,c) \cdot \times G(r,c)$, where R, G, B denotes red, green, and blue channel, '.×' denotes pixelwise-multiplication. After this operation, the blue channel intensity will be increased; therefore, it is normalized to be in the range [0, 255] (using the uint8() function). In this manner, the anterior and posterior spatial information was considered. We used grouped convolutional layers so that the network could compute the separate image. After using three grouped layers, the three obtained images were combined. Behind each grouped convolution layer, a batch normalization layer [15] and rectified linear unit were embedded, which are not displayed in Figure 4.



Figure 4. Pelvis NN for metastasis classification. We applied three NN models for comparisons.

The faster R-CNN will generate six outputs for every detected object: (width, height, center x, center y) of a bounding box, label of class, and confidence of classification. More details can be found in [12].

2.7. Training NN

2.7.1. Hard Negative Mining

Hard negative mining (HNM) is a technique for increasing the specificity performance and was proposed early in the development of computer vision [16,17]. In this method, a model is trained with an initial subset of negative examples. Then, negative examples that are misclassified by this initial model are collected to form another subset of hard negatives. A new model is trained with this new subset, and the aforementioned process may be repeated many times. Our strategy is described in the following text. After the first training, all the training images are fed to the trained network again. A maximum of three false positives (FPs), which have the highest scores for misclassifying metastasis, are collected for each image. All the false-positive boxes are then collected to train the network again.

2.7.2. Hard Positive Mining

A hard positive mining (HPM) approach is proposed in this study to increase the sensitivity performance. This method is also based on an initial trained model. Some positives, which might or might not be correctly detected by the initial model, are collected. We call these positives hard positives because they were not detected or their scores are very low in case they were detected. The HPM technique was implemented as follows. According to an initial trained network, all images are fed into the network to determine if any positive is missing. If a positive is missing, we define its score as 0. Then, all scores of the detected true positives (TPs) are collected to calculate the training number (*tn*), which is defined as follows:

$$tn = \frac{2}{\max(score, 0.1)}.$$

In case a positive is not detected [false negative (FN)], its training number is set to 20. The lower the score of a hard positive is, the higher the training number assigned to it is. In the second training phase, the bounding box of each hard positive is randomly "swung" at its original local area so that the training pattern is never repeated. The aforementioned method is also a type of data augmentation but is more efficient and targeted.

2.8. Performance

In the chest NN, many bounding boxes are output and marked as detected metastasis. These outputs are compared to the boxes marked by physicians manually. We used intersection over union (IoU) 50 as the threshold to define TPs and FPs. If the output box overlaps the physician's manual box by more than 50%, the box is defined as a TP; otherwise, the box is defined as an FP. If a physician's manual box is not detected by any box, the box is defined as an FN. A true negative is not feasible. According to the aforementioned criteria, the precision–recall curve is suitable for determining the performance of networks.

3. Results

The combined structure of all the CNNs adopted in this study is illustrated in Figure 5. All the CNN processes are fully automated. The combined network has two stages. In stage I, a simplified faster R-CNN is used to detect the chest and pelvis area. This R-CNN outputs the bounding boxes. Then, the area in the box is processed, as described in Section 2.6. The colored image is input into the stage II NN. In stage II, two types of NN exist: the chest and pelvis NNs. For the chest NN, we designed a light version of the faster R-CNN, which can be trained in a personal computer with only a single GPU (Nvidia GTX 1080 Ti) card. In the chest NN, we used YOLO v3 as the backbone NN. In the pelvis NN, we compared three NNs: ResNet18, ResNet101, and Inception v3.



Figure 5. Combined structure of all the NNs used in this study. The structure comprises two stages. Scheme 3. Results.

The R-CNN used for the detection of the five body parts could achieve 100% accuracy according to IoU 90. Identifying the five body parts was straightforward because they were significantly different.

The performance of the pelvis NN is presented in Table 2. The 205 patients were divided into ten folds. Table 2 presents the average results of 10-fold training and testing. We controlled the specificity as 0.81 and compared the sensitivities of the aforementioned three CNNs. The results indicate that ResNet101 had the best sensitivity among these three CNNs. Notably, this was the one-time 10-fold cross validation result.

Tab	le	2.	Μ	letastasis	classifi	ication	results	for	the	pelvis	NN.
-----	----	----	---	------------	----------	---------	---------	-----	-----	--------	-----

	Inception v3	ResNet 18	ResNet 101
Sensitivity	0.84 ± 0.13	0.80 ± 0.15	0.87 ± 0.12
Specificity	0.81 ± 0.12	0.81 ± 0.12	0.81 ± 0.11

A total of 85 patients had pelvic bone metastasis, and the remaining 120 patients had no pelvic bone metastasis.

The performance of the chest NN is presented in Table 3. We compared YOLO v3 and the simplified faster R-CNN. YOLO v3 was implemented in the Taiwania II supercomputer (Quanta Computer Inc., Taipei, Taiwan), and the faster R-CNN was implemented in a personal computer with a single Nvidia GTX 1080 Ti (Nvidia, Santa Clara, CA, USA). We used 10-fold training and testing to obtain the average of sensitivity and precision. YOLO v3 had superior performance compared to the faster R-CNN. Moreover, YOLO v3 may have had a deeper network than our simplified faster R-CNN. However, the simplified faster R-CNN could be trained in a personal computer with a single GPU card; however, YOLO v3 could not be trained due to memory limitations. YOLO v3 (Darknet-53) had 53 convolution layers, and the simplified faster R-CNN had only ten deep layers. The qualitative results of chest NN are shown in Figure 6. In this figure, we see the injury (post-trauma lesions) are not detected as metastasis, Figure 6d.

Table 3. Metastasis detection and classification results for the chest NN.

	Yolo v3	Faster R CNN
Sensitivity	0.82 ± 0.08 0.70 ± 0.11	0.70 ± 0.04 0.69 ± 0.07
Trecision	0.70 ± 0.11	0.07 ± 0.07









Figure 6. Detection and classification results of the chest NN. All the red marks denote predicted metastasis lesions by the chest NN (Yolo v3). (a) Case no. 13, TPR = 6/7. (b) Case no. 108, TPR = 8/12. (c) Case no. 127, TPR = 10/10. (d) Case no. 137, FP = 1. (e) Case no. 152, TPR = 8/10. (f) Case no. 165, TPR = 8/9. TPR is true positive rate, FP is false positive.

The learning parameters used in each CNN are listed in Table 4. Except for the faster R-CNN, the other CNNs were executed in the Taiwan computing cloud (TWCC) [18].

Model	Learning Rate	Batch Size	Epoch
Faster R CNN	0.0001	16	40
Yolo v3	0.000579	32	100
ResNet 18	0.001	8	50
ResNet 101	0.001	8	50
Inception v3	0.001	8	50

Table 4. Training parameters of each CNN.

Gradient method: ADAM.

Figure 6 displays the detection and classification results for the chest NN (YOLO). All the red marks were classified as metastasis. Most of the metastasis locations were correctly detected and classified. Some lesions had low luminance in the image; however, they exhibited high luminance in the other side, which is not shown in Figure 6. Figure 6d illustrates four post-trauma lesions, which were correctly detected and classified. The arrow illustrates the positions of four post-trauma lesions. However, there is a false positive.

We found that using hard example mining increased the detection and classification performance. In general, HNM increased the precision and HPM increased the sensitivity. This is a trend that is not guaranteed for every case. Figure 7 illustrates the results obtained with (blue curve) and without (red curve) the use of hard example mining. Based on the result shown in Figure 7, we see a tendency that a superior precision might be achieved when using hard example mining; more experiments can be performed to prove this.



Figure 7. Recall–precision curve. The red curve is obtained without using hard example mining. The blue curve is obtained using hard example mining. The dots are the experimental results, while the curve is the curve-fitting of the dots. Please note that the abscissa is reversed by "1-recall".

4. Discussion

The image normalization process is necessary because it can transform an unclear bone scan raw image into a visible one. We note that many raw images have less intensity, meaning that they are not usable for neural network training purposes. The different contrast levels for data augmentation purposes lead to expanding different mean-intensity levels. This process is, according to our experience, important and helpful, offering more information for the faster R-CNN in recognizing small or unclear lesions. We illustrate some different contrast-level images, as shown in Figure 8. As we increase the intensity, it is important to note that the image should not be oversaturated. Based on our experience, we can control the whole image brightness so that the image will not be oversaturated. Figure 8 from the upper-left to the bottom-right constitutes seven images with different levels of contrast. These images are provided as a simulation while the medical doctor is observing an image. The observer has to change intensity/contrast to see different sites of potential lesions in order to make a correct diagnosis. This augmentation is important in providing different views of the same lesion, especially as some lesions might have strong absorptions and some might not on the same image.



Figure 8. The 7 different contrast-level images were created for data augmentation purpose.

In this study, we proposed the consideration of anterior and posterior images in 3D form to obtain their spatial relation. The third image that should be considered is

a nonlinear combination of the anterior and posterior views. With the aforementioned strategy, metastasis lesions appearing in the front view are red and those appearing in the back view are green. If a lesion has strong absorption of Tc-99m MDP, it would be white. The proposed network can suitably consider the 3D relation because it takes advantage of grouped convolution. In previous studies [6,7], such an arrangement has not been used. Figure 9 provides an example of the 3D formation of a bone scan.



Figure 9. Three-dimensional formation of the anterior and posterior views of a bone scan (the spatial relation is used in the proposed network).

HNM has been used in previous studies; however, HPM has rarely been used. In this study, we used HPM for metastasis detection and classification. This technique provides superior sensitivity in many but not all cases. We believe that HPM is a type of augmentation technique with a targeted purpose.

¹⁸F-Fluoride PET/CT scans (here we shorten it to ¹⁸F-Fluoride scans) represent an alternative way of detecting the bone metastasis of some cancers. ¹⁸F-Fluoride scans can provide 3D information. The maximum intensity projection (MIP) of the volumetric whole-body images from ¹⁸F-Fluoride scans is very similar to planar bone scintigraphy [19]. However, it (MIP) causes similar false positives, such as bone injury and osteophytes, to those that occur in Tc-99m MDP planar bone scintigraphy. Our model cannot be directly applied to the volumetric data provided by ¹⁸F-Fluoride scans, since it is not designed for them. However, our model has the potential to be applied to the MIP of the volumetric data of ¹⁸F-Fluoride scans, using some pre-training techniques such as transfer learning [20].

5. Conclusions

We developed a chest NN and a pelvis NN, which can detect and classify metastasis hotspots. The sensitivity and precision rate for metastasis detection and classification in the chest were 0.82 ± 0.08 and 0.70 ± 0.11 , respectively. The sensitivity and specificity for metastasis classification in the pelvis were 0.87 ± 0.12 and 0.81 ± 0.11 , respectively. The proposed system can be used to obtain a prediagnostic report for physicians' final decisions.

Author Contributions: Data curation, K.-Y.Y.; Funding acquisition, D.-C.C.; Investigation, T.-C.H.; Methodology, D.-C.C.; Project administration, D.-C.C.; Resources, C.-H.K.; Software, D.-C.C. and C.-C.L.; Supervision, D.-C.C. and C.-H.K.; Validation, T.-C.H.; Visualization, T.-C.H.; Writing—original draft, D.-C.C.; Writing—review & editing, D.-C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by MOST, Taiwan, under Grant number MOST 108-2221-E-039-003. This research was partially funded by China Medical University under grant number CMU109-ASIA-02.

Institutional Review Board Statement: This study was approved by the Institutional Review Board of China Medical University and Hospital Research Ethics Committee (CMUH106-REC2-130).

Acknowledgments: We are grateful to the National Center for High-Performance Computing for providing access to its computing facilities.

Conflicts of Interest: The authors have no conflict of interest to declare.

References

- 1. National Health Insurance Research Database. Available online: https://www.mohw.gov.tw/cp-4256-48057-1.html (accessed on 10 April 2019).
- Bubendorf, L.; Schöpfer, A.; Wagner, U. Metastatic patterns of prostate cancer: An autopsy study of 1589 patients. *Hum. Pathol.* 2000, *31*, 578–583. [CrossRef] [PubMed]
- 3. Treating Prostate Cancer Spread to Bones. Available online: https://www.cancer.org/cancer/prostate-cancer/treating/treating-pain.html (accessed on 14 April 2020).
- 4. Imbriaco, M. A new parameter for measuring metastatic bone involvement by prostate cancer: The Bone Scan Index. *Clin. Cancer Res.* **1998**, *4*, 1765–1772. [PubMed]
- Brown, M.S. Computer-Aided Bone Scan Assessment with Automated Lesion Detection and Quantitative Assessment of Bone Disease Burden Changes. U.S. Patent US20140105471, 7 April 2012.
- 6. Ulmert, D. A Novel automated platform for quantifying the extent of skeletal tumour involvement in prostate cancer patients using the bone scan index. *Eur. Urol.* **2012**, *62*, 78–84. [CrossRef] [PubMed]
- 7. Apiparakoon, T.; Rakratchatakul, N.; Chantadisai, M.; Vutrapongwatana, U. MaligNet: Semisupervised learning for bone lesion instance segmentation using bone scintigraphy. *IEEE Access* 2020, *8*, 27047–27066. [CrossRef]
- 8. Sun, P.; Wang, D.; Mok, V.-C.; Shi, L. Comparison of feature selection methods and machine learning classifiers for radiomics analysis in glioma grading. *IEEE Access* **2019**, *7*, 102010–102020. [CrossRef]
- Chen, Y.-H.; Lue, K.-H.; Chu, S.-C. Combing the radiomic features and traditional parameters of ¹⁸F-FDG PET with clinical profiles to improve prognostic stratification in patients with esophageal squamous cell carcinoma treated with neoadjuvant chemo-radiotherapy and surgery. *Ann. Nucl. Med.* 2019, *33*, 657–670. [CrossRef] [PubMed]
- 10. Alom, M.Z. A State-of-the-art Survey on Deep Learning Theory and Architectures. Electronics 2019, 8, 292. [CrossRef]
- Zhao, Z.Q. Object detection with deep learning: A review. *IEEE Trans. Neural Netw. Learn. Syst.* 2019, 30, 3212–3232. [CrossRef] [PubMed]
- 12. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [CrossRef] [PubMed]
- 13. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, NV, USA, 27–30 June 2016.
- 14. PyTorch-YOLOv3. Available online: https://github.com/eriklindernoren/PyTorch-YOLOv3 (accessed on 10 April 2021).
- 15. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv* 2015, arXiv:1502.03167.
- 16. Sung, K.; Poggio, T. *Example-Based Learning for View Based Human Face Detection*; Technical Report A.I. Memo No. 1521; Massachussets Institute of Technology: Cambridge, MA, USA, 1994.
- 17. Felzenszwalb, P.; Girshick, R.; McAllester, D.; Ramanan, D. Object detection with discriminatively trained part based models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1627–1645. [CrossRef] [PubMed]
- 18. Taiwan Computing Cloud. Available online: https://www.twcc.ai/ (accessed on 10 April 2021).
- Löfgren, J.; Mortensen, J.; Rasmussen, S.H.; Madsen, C.; Loft, A.; Hansen, A.E.; Oturai, P.; Jensen, K.E.; Mørk, M.L.; Reichkendler, M.; et al. A Prospective Study Com- paring 99mTc-Hydroxyethylene-Diphosphonate Planar Bone Scintigra- phy and Whole-Body SPECT/CT with 18F-Fluoride PET/CT and 18F-Flu- oride PET/MRI for Diagnosing Bone Metastases. *J. Nucl. Med.* 2017, 58, 1778–1785. [CrossRef] [PubMed]
- 20. Pratt, L.Y.; Thrun, S. Machine Learning—Special Issue on Inductive Transfer; Springer: Berlin/Heidelberg, Germany, 1997.