

## Article

# Deep-Learning-Based Detection of Cranio-Spinal Differences between Skeletal Classification Using Cephalometric Radiography

Seung Hyun Jeong <sup>1,†</sup>, Jong Pil Yun <sup>1,†</sup>, Han-Gyeol Yeom <sup>2</sup>, Hwi Kang Kim <sup>3</sup> and Bong Chul Kim <sup>3,\*</sup> 

<sup>1</sup> Safety System Research Group, Korea Institute of Industrial Technology (KITECH), Gyeongsan 38408, Korea; shjeong@kitech.re.kr (S.H.J.); rebirth@kitech.re.kr (J.P.Y.)

<sup>2</sup> Department of Oral and Maxillofacial Radiology, Daejeon Dental Hospital, Wonkwang University College of Dentistry, Daejeon 35233, Korea; hangyeol1214@gmail.com

<sup>3</sup> Department of Oral and Maxillofacial Surgery, Daejeon Dental Hospital, Wonkwang University College of Dentistry, Daejeon 35233, Korea; hwi1304@naver.com

\* Correspondence: bck@wku.ac.kr

† S.H.J. and J.P.Y. contributed equally to this study.

**Abstract:** The aim of this study was to reveal cranio-spinal differences between skeletal classification using convolutional neural networks (CNNs). Transverse and longitudinal cephalometric images of 832 patients were used for training and testing of CNNs (365 males and 467 females). Labeling was performed such that the jawbone was sufficiently masked, while the parts other than the jawbone were minimally masked. DenseNet was used as the feature extractor. Five random sampling crossvalidations were performed for two datasets. The average and maximum accuracy of the five crossvalidations were 90.43% and 92.54% for test 1 (evaluation of the entire posterior–anterior (PA) and lateral cephalometric images) and 88.17% and 88.70% for test 2 (evaluation of the PA and lateral cephalometric images obscuring the mandible). In this study, we found that even when jawbones of class I (normal mandible), class II (retrognathism), and class III (prognathism) are masked, their identification is possible through deep learning applied only in the cranio-spinal area. This suggests that cranio-spinal differences between each class exist.

**Keywords:** machine learning; artificial intelligence; malocclusion; diagnostic imaging



**Citation:** Jeong, S.H.; Yun, J.P.; Yeom, H.-G.; Kim, H.K.; Kim, B.C. Deep-Learning-Based Detection of Cranio-Spinal Differences between Skeletal Classification Using Cephalometric Radiography. *Diagnostics* **2021**, *11*, 591. <https://doi.org/10.3390/diagnostics11040591>

Academic Editor: Sameer Antani

Received: 15 February 2021

Accepted: 22 March 2021

Published: 25 March 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/>).

## 1. Introduction

Dentofacial dysmorphism exhibits various aspects such as prognathism, retrognathism, maxillary hypoplasia, and asymmetry [1,2]. For their treatment, several techniques of orthognathic surgery or orthodontics are applied [2–4]. Meanwhile, the stomatognathic system is composed of static and dynamic structures, and its harmonious functioning is based on the balanced relationship between them [5]. In addition, hard and soft cephalic structures arise, grow, and organize in a mutual balance [6]. Cranio-facial skeletons constantly reflect these influences and their related functional conditions [1,6,7]. Therefore, the genesis of a malocclusion is usually linked to an impairment of some kind to eugnathic growth that involves to various extents the mandible, the maxilla, and the functional matrix (tongue and facial muscles) [5].

Until now, orthodontics and orthognathic surgery have mainly relied on linear and angular measurements for the diagnosis and the planning of the therapeutic procedures [1,3,7–13]. These measurements depend on the identification of several landmarks on cephalometric images, which are then applied to define the aforementioned measurements [1,3,7–13]. It is well recognized that the relation between these metrics varies with the type of bite and therefore is different in skeletal classes I, II, and III [1,7,13]. In addition, most of these landmarks on cephalometric images are concentrated in the maxilla and mandible [7]. However, the authors wondered if the difference between skeletal classes I, II, and III

appears only in maxilla and mandible or if not, if it is also revealed in the cranio-spinal area excluding jaw. We also wanted to find a way to intuitively distinguish skeletal classes I, II, and III without linear and angular measurements.

Advances in convolutional neural networks (CNNs) continue [14–17]. They are being applied in a variety of dental and maxillofacial fields. For instance, they are used to assess soft-tissue profiling and extraction difficulty for mandibular third molars [18,19]. In addition, Xiao et al. proposed an end-to-end deep-learning framework to estimate patient-specific reference bony shape models for patients with orthognathic deformities [20]. Moreover, Sin et al. evaluated an automatic segmentation algorithm for pharyngeal airway in cone-beam-computed tomography images [21]. CNNs have proven their applicability to dental and maxillofacial fields through many other studies. However, to the best of our knowledge, CNNs have not been applied yet to clarify cranio-spinal differences between skeletal classification. Therefore, the aim of this study was to reveal cranio-spinal differences between skeletal classification using CNNs.

## 2. Materials and Methods

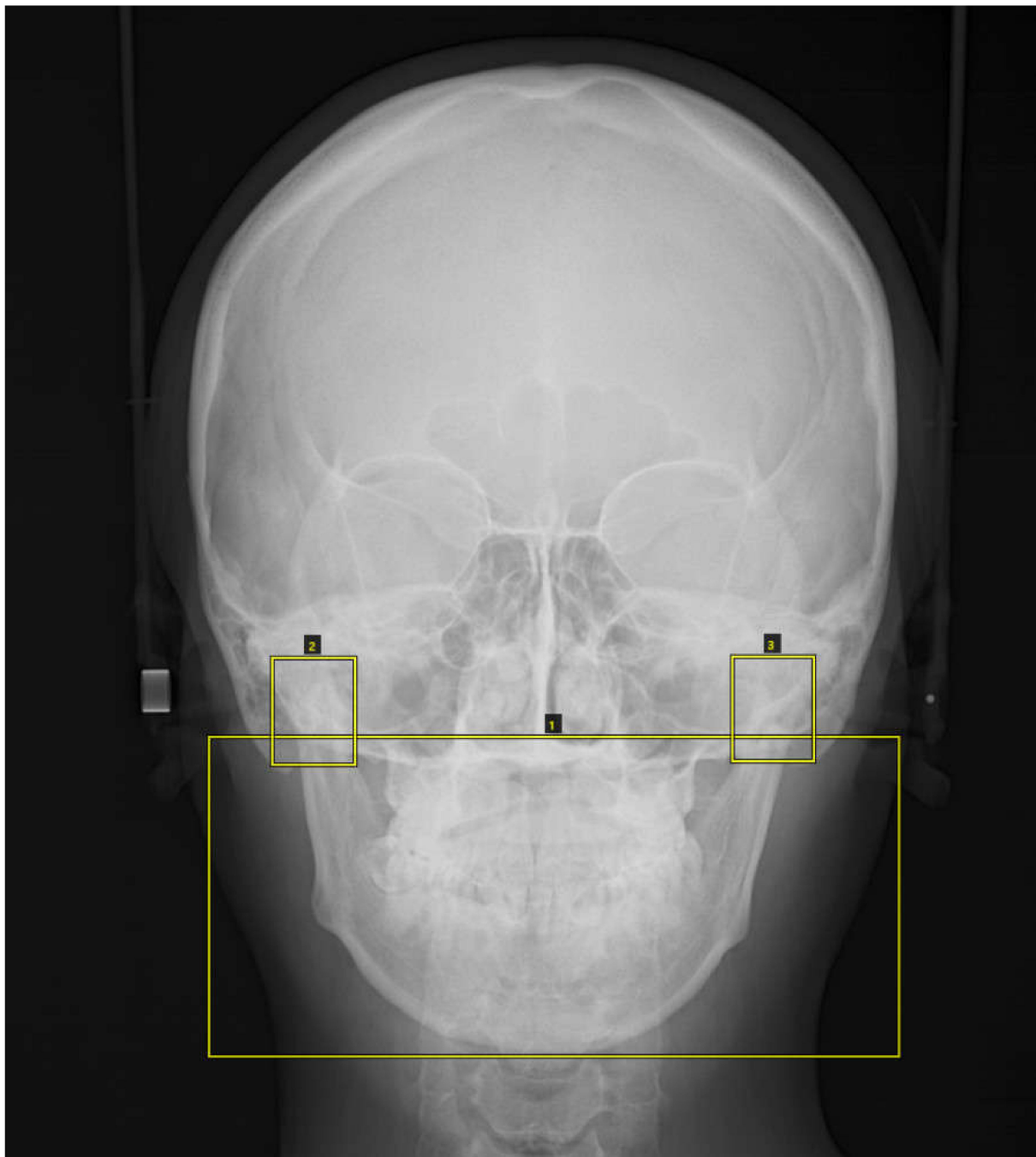
### 2.1. Datasets

In this study, transverse and longitudinal cephalometric images of 832 Korean patients who visited Daejeon Dental Hospital, Wonkwang University between January 2007 and December 2019 complaining about dentofacial dysmorphism and/or a malocclusion were used for the training and testing of a deep-learning model (365 males and 467 females with a mean age of  $18.37 \pm 8.06$  years). Patients with a congenital deformity, infection, trauma, or tumor history were excluded. The lateral and posterior–anterior (PA) cephalometric images were obtained using a Planmeca Promax (Planmeca OY, Helsinki, Finland), and the images were extracted in JPG format. The original images had a pixel resolution of  $2045 \times 1816$  with a size of 0.132 mm/pixel.

All radiographic images were annotated by two orthodontists, two oral and maxillofacial surgeons, and one oral and maxillofacial radiologist. Point A–nasion–point B (ANB) and a Wits appraisal were used to diagnose the sagittal skeletal relationship. Jarabak's ratio and Björk's sum were used to determine the vertical skeletal relationship. With consensus of five specialists, patients' skeletal type was determined: class I ( $n = 272$ , 111 males and 161 females with a mean age of  $17.17 \pm 8.28$  years); class II ( $n = 294$ , 105 males and 189 females with a mean age of  $19.47 \pm 8.85$  years); or class III ( $n = 266$ , 149 males and 117 females with a mean age of  $18.36 \pm 6.61$  years).

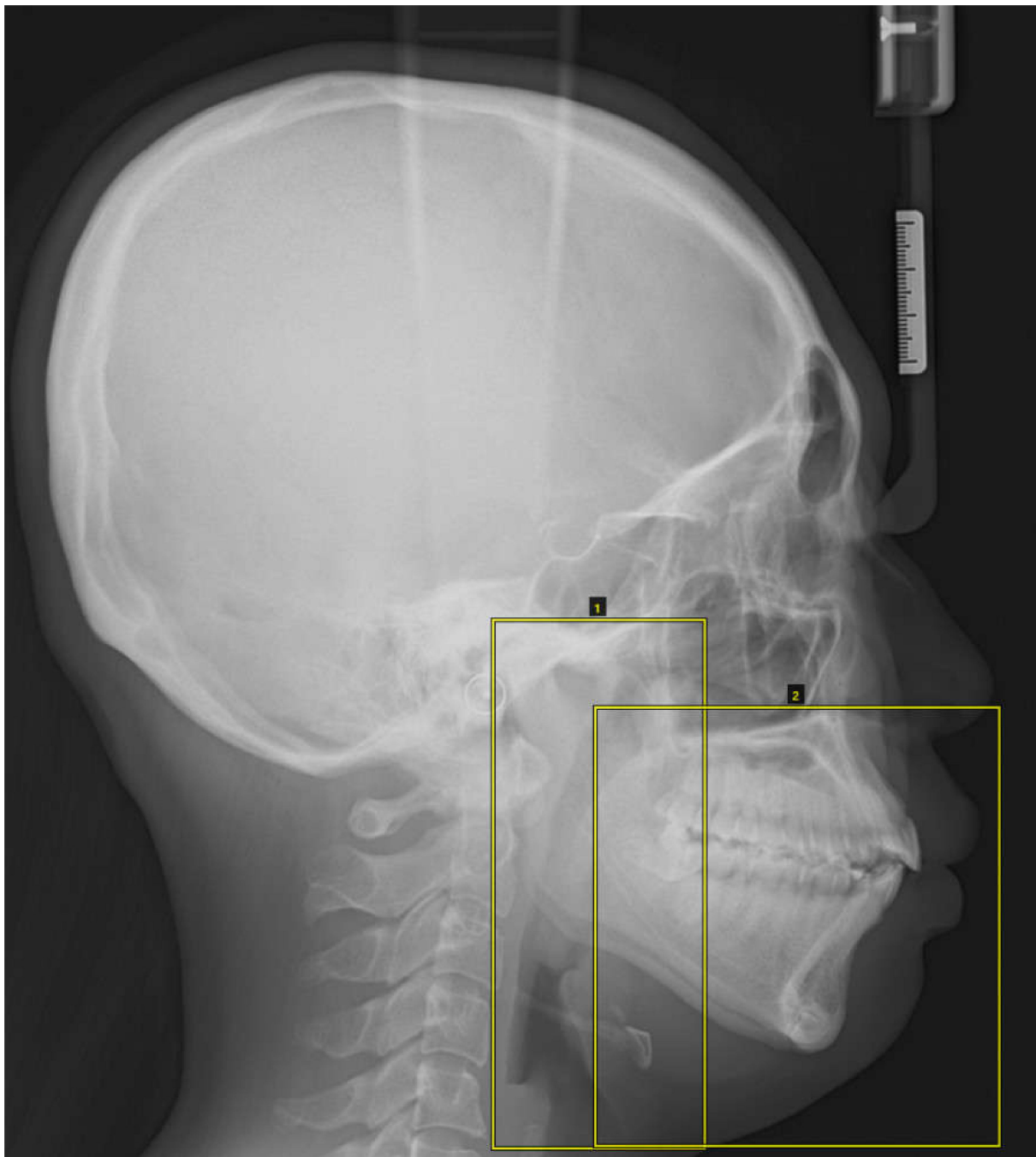
The purpose of this study was to determine whether there is an additional structural difference that makes it possible to distinguish the skeletal class in the structures of the head and neck other than the jawbone. Thus, labeling was manually performed such that the jawbone was sufficiently masked while the parts other than the jawbone were minimally masked.

The PA cephalometric images were masked with three square markers: a lower large square containing maxilla and mandible (nasal floor and hard palate region~inferior border of mandible) plus right and left small squares containing the condylar process (Figure 1).



**Figure 1.** The posterior–anterior (PA) cephalometric images were masked with three square markers: a lower large square containing maxilla and mandible (nasal floor and hard palate region ~ inferior border of mandible) plus right and left small squares containing the condylar process.

The lateral cephalometric images were labeled with two square markers: a left long square containing the condylar process, coronoid process, mandibular ramus, and airway space and a right square containing the dentoalveolar region, maxilla, mandibular body, and lower facial soft tissue (Figure 2).



**Figure 2.** The lateral cephalometric images were masked with two square markers: a left long square containing the condylar process, coronoid process, mandibular ramus, and airway space and a right square containing the dentoalveolar region, maxilla, mandibular body, and lower facial soft tissue.

## 2.2. Preprocessing and Image Augmentation

Each patient's PA and lateral cephalometric images were preprocessed for training. The acquired data were resized to have the same height and width for training because they were different for each patient. For image resizing, we used OpenCV's API based on interpolation. Given that the skeleton is classified according to a geometric relationship, the height and width were resized at the same ratio. The height of the original cephalometric images was resized to 500, and the corresponding height ratio was applied to the width. After that, the cephalometric images were placed at the middle and zero-padding was performed to obtain  $500 \times 500$  images. Note that the masking process mentioned in the

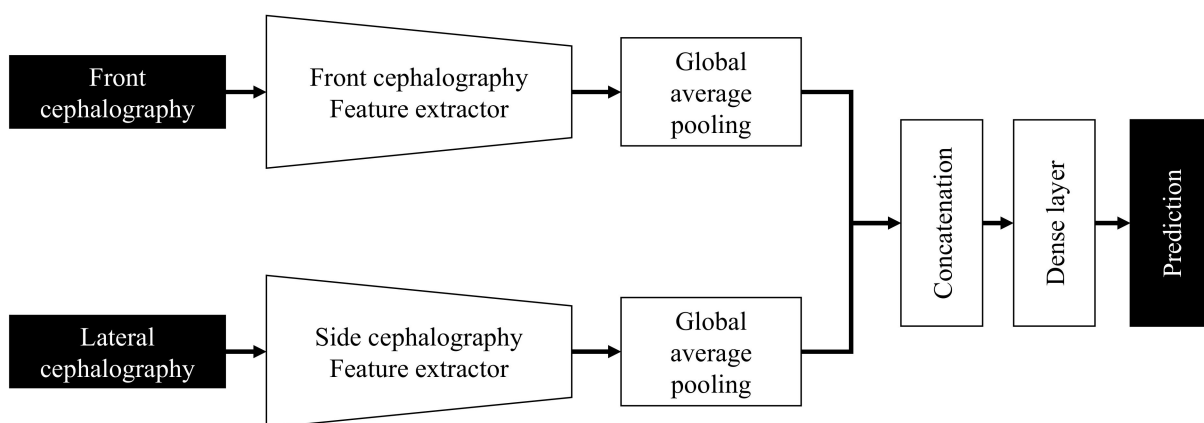
dataset paragraph was applied after image resizing. In addition, data augmentation was performed for the preprocessed images to improve accuracy and prevent overfitting by using Pytorch's color jitter and random horizontal flip. Finally, the data were normalized by using the following equation.

$$p_{i,j}^* = \frac{p_{i,j} - \text{mean}}{\text{std.}} \quad (1)$$

where  $p_{i,j}^*$ ,  $p_{i,j}$ , mean, and std. are normalized pixel value, original pixel value, mean, and standard deviation, respectively. The values of mean and std. are set to 0.5 and 0.5, respectively.

### 2.3. Architecture of the Deep CNN

The network structure that classifies into skeletal classes I, II, and III using PA and lateral cephalometric images is shown in Figure 3. The PA and lateral cephalometric images were fed to each feature extractor to obtain the feature map separately. Various backbone networks, such as VGG [22], ResNet [23], and DenseNet [24], can be used as feature extractors, and feature maps of different dimensions can be obtained according to each network's structure. In this study, DenseNet, proposed in 2016, was used as a feature extractor. DenseNet is a network that extracts features by continuously connecting the feature map of the previous layers with the input of the next layer. Figure 4 shows a five-layer dense block with a growth rate  $k = 4$ . ResNet is a method consisting in adding feature maps, while DenseNet is a structure that concatenates feature maps. Through this structure, the vanishing gradient can be improved, and the feature propagation can be reinforced. The depth of the feature map extracted through DenseNet is determined according to the growth rate and the number of layers of each block, whereas the width and height are determined according to the number of downsamplings. In this study, because pretrained DenseNet121 was used, an input image of  $500 \times 500 \times 3$  is converted into a feature map of  $15 \times 15 \times 1024$  after passing through the feature extractor. Each feature map output from the PA and lateral cephalometric images is transformed into a vector through global average pooling and merged into one vector through concatenation. The final classification is performed through a dense layer. The proposed network was implemented using Pytorch 1.2.



**Figure 3.** Multiside convolutional neural networks (CNNs) for classification using PA and lateral cephalometric images.

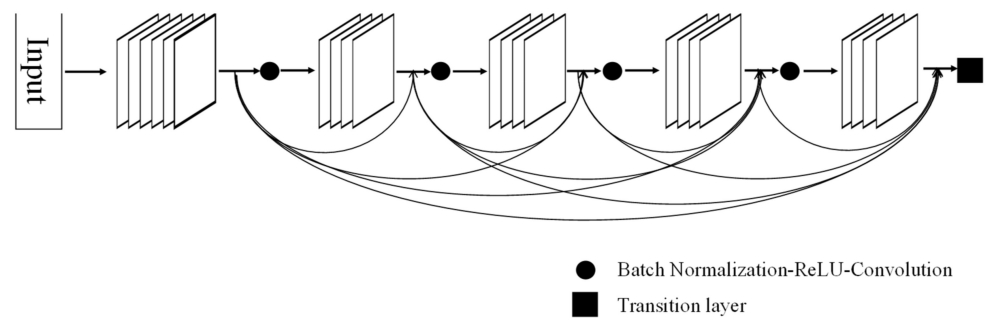


Figure 4. A five-layer dense block with a growth rate  $k = 4$ .

2.4. Visualization Method

In this study, the feature map was displayed so that the part extracted as a feature of the PA and lateral cephalometric images could be confirmed. The class activation map proposed in 2015 was used as a display method [25]. The class activation map is calculated as the summation of each feature map multiplied by the corresponding weight value of the dense layer, as shown in Figure 5. Through this method, it is possible to check which part of the cephalometric image was activated for classification. The greater the activation, the redder it is; and the greater the inactivation, the bluer it is.

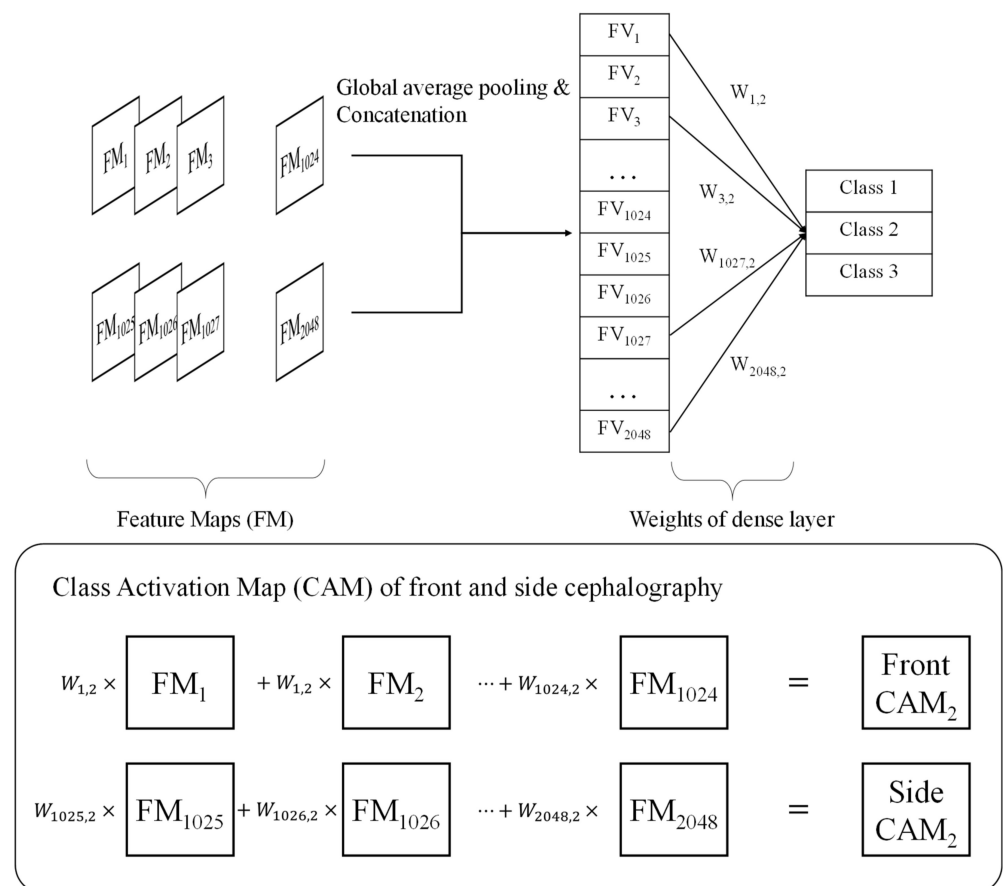
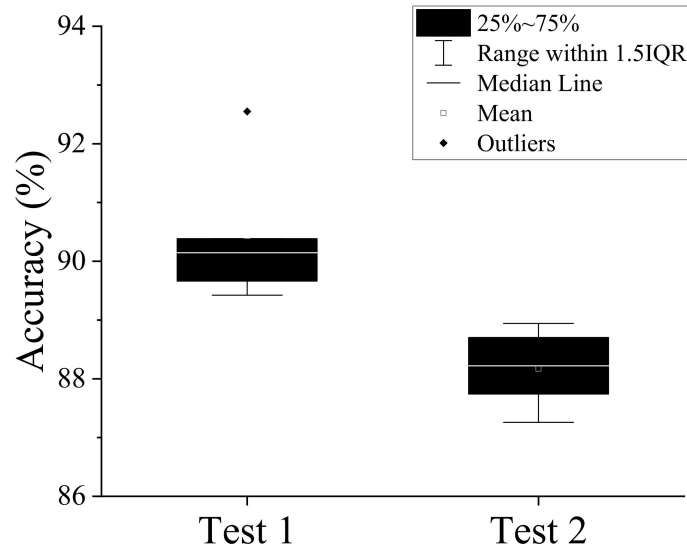


Figure 5. Class activation map (CAM) generation of PA and lateral cephalometric images.

3. Results

The proposed CNNs were trained using the Adam optimizer [26]. The initial learning rate was set to 0.001. The learning rate decay was set to 0.95 and was applied every five epochs. To take into account the randomness of the deep-learning-network training

algorithm, five random sampling crossvalidations were performed for two datasets. The average and maximum accuracy of the five crossvalidations were 90.43% and 92.54% for test 1 (evaluation of the entire posterior–anterior (PA) and lateral cephalometric images) and 88.17% and 88.70% for test 2 (evaluation of the PA and lateral cephalometric images obscuring the mandible). A box plot of the accuracy for each test is shown in Figure 6. Table 1 shows the confusion matrix for best accuracy result of each test.



**Figure 6.** Test accuracies in five random sampling crossvalidations. Test 1: evaluation of the entire PA and lateral cephalometric images. Test 2: evaluation of the PA and lateral cephalometric images obscuring the mandible.

**Table 1.** Confusion matrices of best accuracy results for (a) test 1 and (b) test 2.

		(a)		
		Predictions		
		Class I	Class II	Class III
Ground Truth	Class I	125	9	7
	Class II	11	141	0
	Class III	3	1	119
		(b)		
		Predictions		
		Class I	Class II	Class III
Ground Truth	Class I	118	12	8
	Class II	17	136	0
	Class III	8	2	115

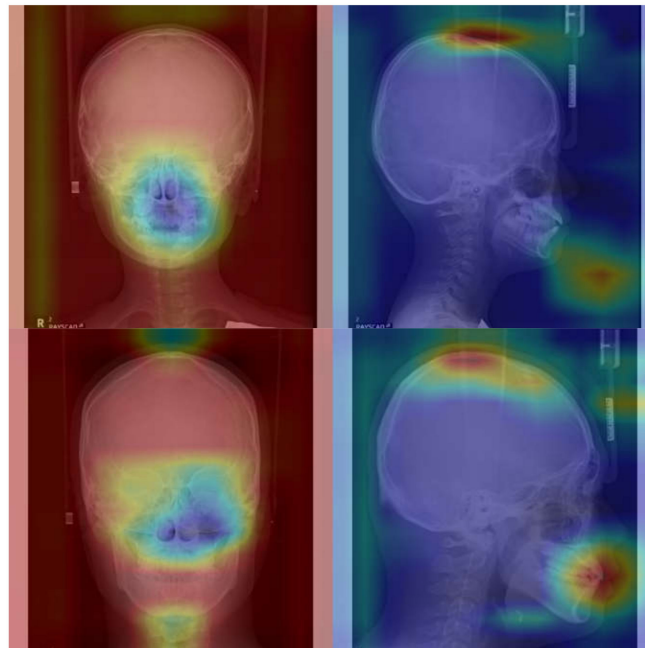
Class I: normal mandible; Class II: retrognathism; and Class III: prognathism. Ground truth: actual group of patients classified according to their mandibular class. Prediction: mandibular class predicted by deep learning. Test 1: evaluation of the entire PA and lateral cephalometric images. Test 2: evaluation of the PA and lateral cephalometric images obscuring the mandible.

#### 4. Discussion

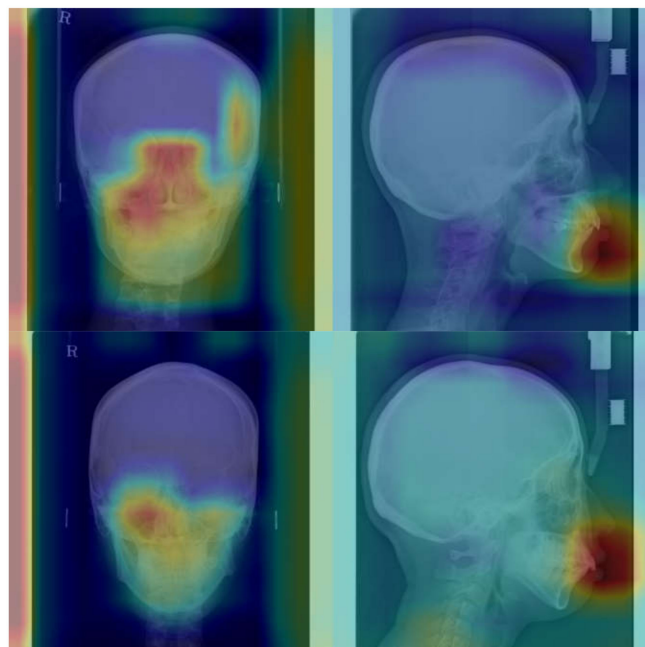
The average and maximum accuracies of the five crossvalidations were 90.43% and 92.54% for test 1, and 88.17% and 88.70% for test 2. As expected, the predicted results were more accurate in test 1, in which all cephalometric images could be analyzed without masking the jawbone. However, the difference in accuracy between test 1 and test 2 was within 5%, which is not significant.

At the same time, with the class activation map, it is possible to know where the CNNs focused on the cephalometric images to provide a prediction (Figure 7). As might be

expected, test 1 focused on the jawbone, especially on the state of the dentition. However, in test 2, the jawbone and dentition were obscured and could not be analyzed. Therefore, CNNs were forced to identify the remaining uncovered regions, that is, the cranio-spinal area. Figure 7 shows the wide area of the cranio-spinal area excluding the jawbone, which is hidden, marked in red. This reveals that the cranio-spinal area is discernibly different in skeletal classes I, II, and III.



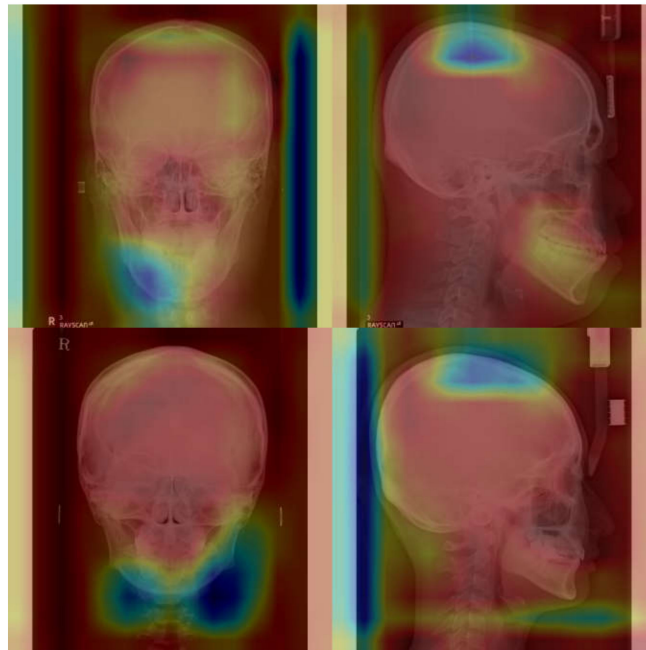
(a)



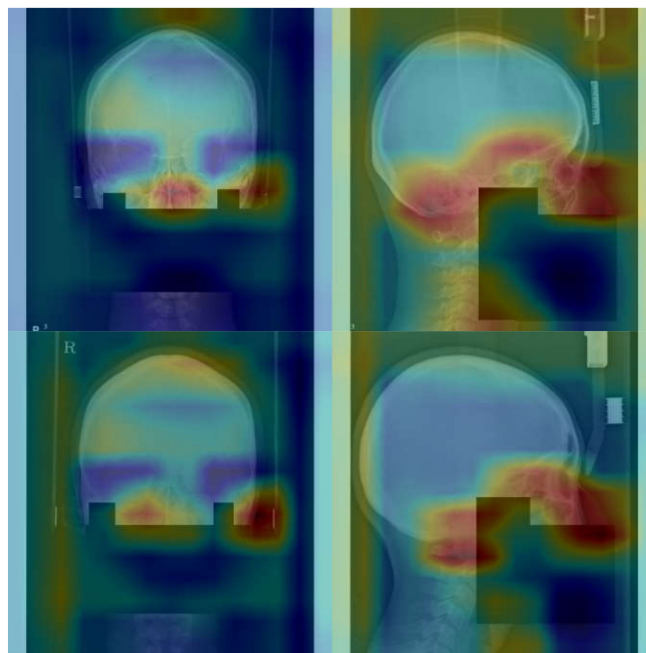
(b)

**Figure 7.** *Cont.*



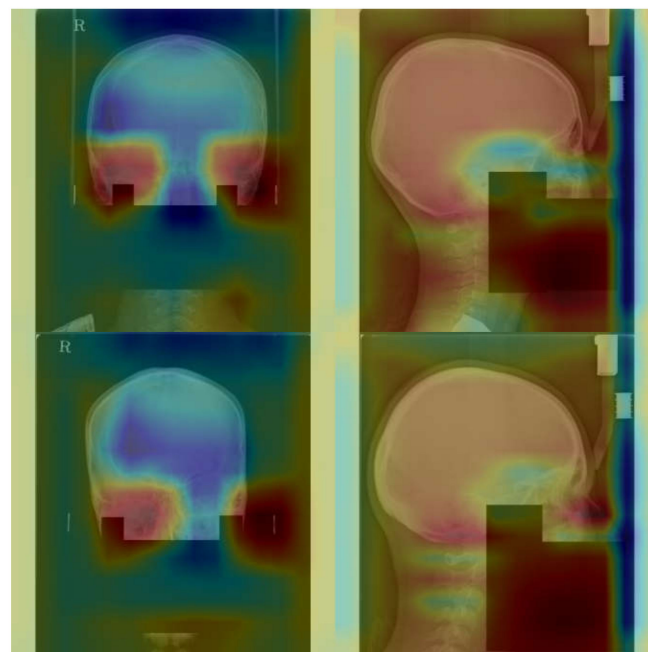


(c)

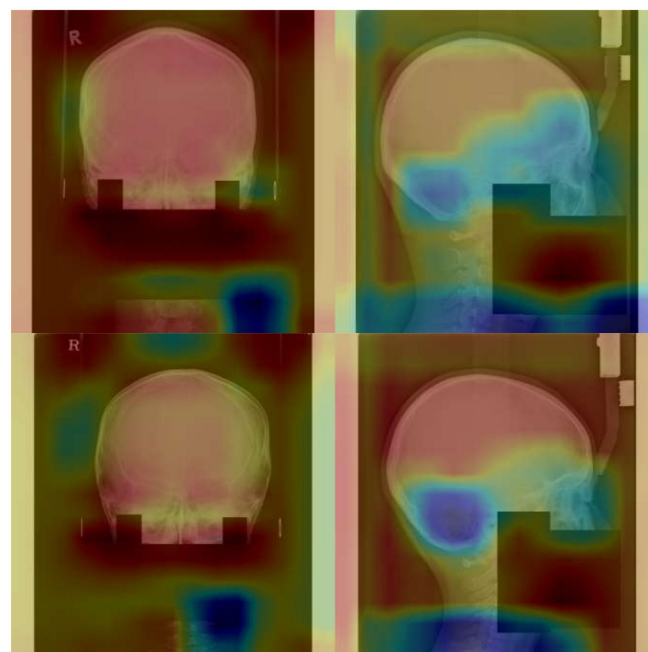


(d)

Figure 7. Cont.



(e)



(f)

**Figure 7.** CAM of PA and lateral cephalometric images for (a) Class I of test 1, (b) Class II of test 1, (c) Class III of test 1, (d) Class I of test 2, (e) Class II of test 2, and (f) Class III of test 2. Class I: normal mandible; Class II: retrognathism; and Class III: prognathism. The greater the activation, the redder it is; the greater inactivation, the bluer it is.

## 5. Conclusions

In this study, we found that even when the jawbones of skeletal classes I, II, and III are masked, their identification is possible through deep learning applied only in the cranio-spinal area. This suggests that cranio-spinal differences exist between each class. Further research is required about where and how cranio-spinal differences emerge.

**Author Contributions:** The study was conceived by B.C.K., who also setup the experiments. S.H.J. and J.P.Y. performed the experiments. H.-G.Y., H.K.K., and B.C.K. generated the data. All authors analyzed and interpreted the data. S.H.J., J.P.Y., and B.C.K. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by a grant from National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) (No. 2020R1A2C1003792).

**Institutional Review Board Statement:** This study was performed in accordance with the guidelines of the World Medical Association Helsinki Declaration for Biomedical Research Involving Human Subjects and was approved by the Institutional Review Board of Daejeon Dental Hospital, Wonkwang University (W2007/004-001, July 8 2020).

**Informed Consent Statement:** Patient consent was waived by the IRBs because of the retrospective nature of this investigation and the use of anonymized patient data.

**Data Availability Statement:** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request but is subject to the permission of the Institutional Review Boards of the participating institutions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Mun, S.H.; Park, M.; Lee, J.; Lim, H.J.; Kim, B.C. Volumetric characteristics of prognathic mandible revealed by skeletal unit analysis. *Ann. Anat. Anat. Anz.* **2019**, *226*, 3–9. [[CrossRef](#)] [[PubMed](#)]
- Lanteri, V.; Cavagnetto, D.; Abate, A.; Mainardi, E.; Gaffuri, F.; Ugolini, A.; Maspero, C. Buccal bone changes around first permanent molars and second primary molars after maxillary expansion with a low compliance ni-ti leaf spring expander. *Int. J. Environ. Res. Public Health* **2020**, *17*, 9104. [[CrossRef](#)]
- Park, J.C.; Lee, J.; Lim, H.J.; Kim, B.C. Rotation tendency of the posteriorly displaced proximal segment after vertical ramus osteotomy. *J. Cranio-Maxillo-Facial Surg.* **2018**, *46*, 2096–2102. [[CrossRef](#)]
- Abate, A.; Cavagnetto, D.; Fama, A.; Matarese, M.; Lucarelli, D.; Assandri, F. Short term effects of rapid maxillary expansion on breathing function assessed with spirometry: A case-control study. *Saudi Dent. J.* **2020**. [[CrossRef](#)]
- Abate, A.; Cavagnetto, D.; Fama, A.; Maspero, C.; Farronato, G. Relationship between breastfeeding and malocclusion: A systematic review of the literature. *Nutrients* **2020**, *12*, 3688. [[CrossRef](#)] [[PubMed](#)]
- Delaire, J.; Schendel, S.A.; Tulasne, J.F. An architectural and structural craniofacial analysis: A new lateral cephalometric analysis. *Oral Surg. Oral Med. Oral Pathol.* **1981**, *52*, 226–238. [[CrossRef](#)]
- Lee, S.H.; Kil, T.J.; Park, K.R.; Kim, B.C.; Kim, J.G.; Piao, Z.; Corre, P. Three-dimensional architectural and structural analysis—a transition in concept and design from delaire’s cephalometric analysis. *Int. J. Oral Maxillofac. Surg.* **2014**, *43*, 1154–1160. [[CrossRef](#)] [[PubMed](#)]
- Shin, H.; Park, M.; Chae, J.M.; Lee, J.; Lim, H.J.; Kim, B.C. Factors affecting forced eruption duration of impacted and labially displaced canines. *Am. J. Orthod. Dentofac. Orthop.* **2019**, *156*, 808–817. [[CrossRef](#)] [[PubMed](#)]
- Kim, B.C.; Bertin, H.; Kim, H.J.; Kang, S.H.; Mercier, J.; Perrin, J.P.; Corre, P.; Lee, S.H. Structural comparison of hemifacial microsomia mandible in different age groups by three-dimensional skeletal unit analysis. *J. Cranio-Maxillo-Facial Surg.* **2018**, *46*, 1875–1882. [[CrossRef](#)]
- Kim, H.J.; Kim, B.C.; Kim, J.G.; Zhengguo, P.; Kang, S.H.; Lee, S.H. Construction and validation of the midsagittal reference plane based on the skull base symmetry for three-dimensional cephalometric craniofacial analysis. *J. Craniofacial Surg.* **2014**, *25*, 338–342. [[CrossRef](#)]
- Kim, B.C.; Lee, S.H.; Park, K.R.; Jung, Y.S.; Yi, C.K. Reconstruction of the premaxilla by segmental distraction osteogenesis for maxillary retrusion in cleft lip and palate. *Cleft Palate-Craniofacial J.* **2014**, *51*, 240–245. [[CrossRef](#)]
- Kang, Y.H.; Kim, B.C.; Park, K.R.; Yon, J.Y.; Kim, H.J.; Tak, H.J.; Piao, Z.; Kim, M.K.; Lee, S.H. Visual pathway-related horizontal reference plane for three-dimensional craniofacial analysis. *Orthod. Craniofacial Res.* **2012**, *15*, 245–254. [[CrossRef](#)] [[PubMed](#)]
- Park, W.; Kim, B.C.; Yu, H.S.; Yi, C.K.; Lee, S.H. Architectural characteristics of the normal and deformity mandible revealed by three-dimensional functional unit analysis. *Clin. Oral Investig.* **2010**, *14*, 691–698. [[CrossRef](#)]
- Awan, M.J.; Rahim, M.S.M.; Salim, N.; Mohammed, M.A.; Garcia-Zapirain, B.; Abdulkareem, K.H. Efficient detection of knee anterior cruciate ligament from magnetic resonance imaging using deep learning approach. *Diagnostics* **2021**, *11*. [[CrossRef](#)]
- Jeon, Y.; Lee, K.; Sunwoo, L.; Choi, D.; Oh, D.Y.; Lee, K.J.; Kim, Y.; Kim, J.W.; Cho, S.J.; Baik, S.H.; et al. Deep learning for diagnosis of paranasal sinusitis using multi-view radiographs. *Diagnostics* **2021**, *11*, 250. [[CrossRef](#)]
- Kumar Singh, V.; Abdel-Nasser, M.; Pandey, N.; Puig, D. Lunginseg: Segmenting covid-19 infected regions in lung ct images based on a receptive-field-aware deep learning framework. *Diagnostics* **2021**, *11*, 158. [[CrossRef](#)] [[PubMed](#)]

17. Singh, G.; Al'Aref, S.J.; Lee, B.C.; Lee, J.K.; Tan, S.Y.; Lin, F.Y.; Chang, H.J.; Shaw, L.J.; Baskaran, L.; On Behalf Of The, C.; et al. End-to-end, pixel-wise vessel-specific coronary and aortic calcium detection and scoring using deep learning. *Diagnostics* **2021**, *11*, 215. [[CrossRef](#)] [[PubMed](#)]
18. Jeong, S.H.; Yun, J.P.; Yeom, H.G.; Lim, H.J.; Lee, J.; Kim, B.C. Deep learning based discrimination of soft tissue profiles requiring orthognathic surgery by facial photographs. *Sci. Rep.* **2020**, *10*, 16235. [[CrossRef](#)] [[PubMed](#)]
19. Yoo, J.H.; Yeom, H.G.; Shin, W.; Yun, J.P.; Lee, J.H.; Jeong, S.H.; Lim, H.J.; Lee, J.; Kim, B.C. Deep learning based prediction of extraction difficulty for mandibular third molars. *Sci. Rep.* **2021**, *11*, 1954. [[CrossRef](#)]
20. Xiao, D.; Lian, C.; Deng, H.; Kuang, T.; Liu, Q.; Ma, L.; Kim, D.; Lang, Y.; Chen, X.; Gateno, J.; et al. Estimating reference bony shape models for orthognathic surgical planning using 3d point-cloud deep learning. *IEEE J. Biomed. Health Inform.* **2021**. [[CrossRef](#)]
21. Sin, Ç.; Akkaya, N.; Aksoy, S.; Orhan, K.; Öz, U. A deep learning algorithm proposal to automatic pharyngeal airway detection and segmentation on cbct images. *Orthod. Craniofacial Res.* **2021**. [[CrossRef](#)] [[PubMed](#)]
22. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the The International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
23. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
24. Huang, G.; Liu, Z.; Maaten, L.v.d.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
25. Zhou, B.; Khosla, A.; Lapedriza, A.; Oliva, A.; Torralba, A. Learning deep features for discriminative localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.
26. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.