Project Name: **CNN** in Bone Age recognition
Name: Alec Mabhiza Chirawu
Chinese Name: 亚历克上
Student Number: M202161029

Neural Networks and Deep Learning
Information And Engineering Department
USTB

Professor: 黄旗明(Roland)

22 April 2022

Summary

Bone age recognition of live humans and bodies is becoming increasingly popular in forensic medicine, choosing athletes and criminal identification. Manual reading methods involve a significant amount of time and effort. To do pre-processing, we use a data set of X-ray scanned pictures of the human left hand. Bone age recognition is a method of determining the age of a person's bones. The focus has shifted away from X-rays and toward MRI-based imaging and using deep active learning technology, we have created a flexible framework for medical imaging and analysis. The difficulty of training a deep neural network(DNN) and Convolutional Neural Network(CNN) with annotated training dataset is addressed.

TABLE OF CONTENTS

1 Introduction

The use of convolution has been one of the most important repercussions of deep learning since 2012. Using neural networks, considerable advances in object detection and recognition can be made. Each neuron in every network we've seen so far is connected to the first hidden layer, which includes all neurons in the input layer. If the input has a lot of dimensions, this won't function. This is due to the fact that each neuron has several connections. Each neuron, for example, has 10,000 parameters if the input is a small 100x100 pixel image (that is, the input vector has a dimension of 10,000)[5]. Choose This is more efficient and can force each neuron to connect to the input in a limited number of ways.The connection arrangement can be tailored to the input's structure. In the case of pictures, for example, the connection patterns imply that neurons can only look at pixels in the input image that are next to them. This concept can be expanded to compel local connectivity at multiple layers, resulting in a deep locally linked network. Gradient descent training is possible because the back propagation algorithm can be modified to deal with local connectivity: in the forward pass, we can compute the values of neurons by assuming that the empty connections have weights of zeros; however, we do not need to compute the gradients for the empty connections in the backward pass.Local networks, locally connected networks, and local receptive field networks are all terms used to describe this type of network. The last name is derived from the fact that neurons in the brain are usually coupled locally, and the term "local receptive field" is used in neuroscience and biology to describe this. The number of connections is greatly reduced when locality structures are used. Another approach known as "weight distribution" can help you save even more money. Some model parameters are forced to be equal because of the weight distribution. In the following layer of the network, for example, w1 = w6, w2 = w4, and w3 = w5. Because of these constraints, the model is quite small in terms of numbers.
The real numbers. Save only a couple of the weights instead of all of them. The principle of sharing the weights is similar to convolution, a common signal processing method. We can apply a "filter" a set of weights to several places in the input signals using convolution. This type of network also has a layer known as the "maxpooling layer" in practise. The maxpooling layer computes the maximum value of a subset of convolutional layer output neurons and utilizes it as an input to higher layers. Convolutional neural networks are another name for this sort of network sometimes called convnets.The sub sampling layer is another name for the CNN maxpooling layer[5,10].

This is due to the fact that the input data is much smaller.We must also mention the rectifier, which is an activation feature specified as $f(x) = max(0, x)$ in the context of synthetic neural networks, where x represents the neuron's input. This activation feature is thought to be more biologically feasible than the commonly utilized logistic sigmoid that is stimulated via way of means of possibility theory. As of 2015, the rectifier is the most well-known activation feature for deep neural networks. A rectified linear unit is another term for a unit that uses a rectifier (ReLU).

1.1 Methodology

The CNN model is most commonly used for segmentation and classification of bone age. Among all the deep learning models, CNN gives effective performance for image segmentation, prediction, and classification. Two-dimensional CNNs (2D-CNNs) and 3D-CNNs.

2. Findings

2.1 Off the shelf CNN feature extraction

X-ray images are markedly different from the real world images commonly employed for training general-purpose CNNs. In this case, it is a convenient alternative to train the regression network, which takes as input, features extracted from one of the deepest layers of off-the-shelf CNN models. Our architecture consists of a pre-trained CNN with the final layers removed so that the output of a fully-connected layer is exposed.

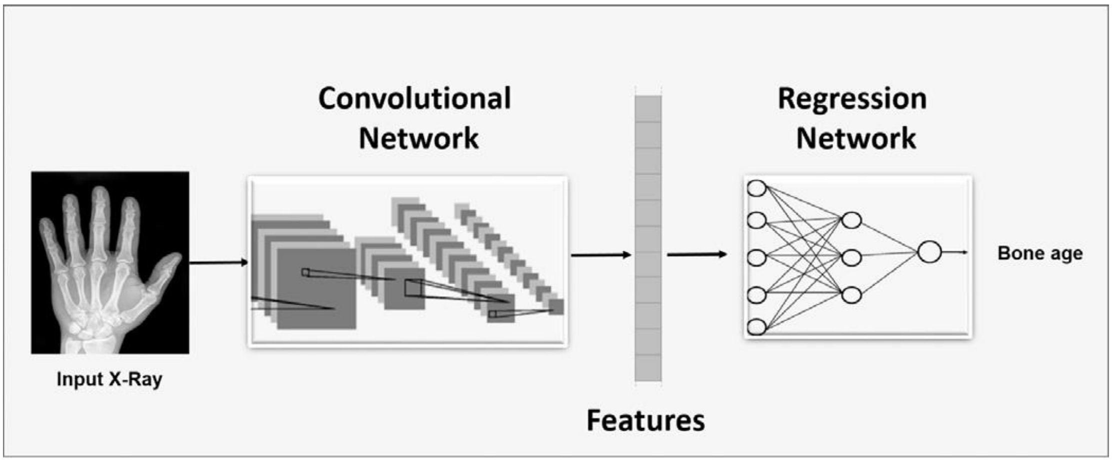

Figure 1: General architecture of deep learning methods for bone age assessment.
It consists of a) a convolutional network consisting of an arbitrary number of convolutional layers (that can be derived from pre-trained CNN models or can be designed from scratch) for feature extraction; and b) a regression network consisting of a set of fully connected layers (generally one or two) and a linear scalar output layer providing the bone age estimate.

2.2 Fine-tuning CNNs trained on general imagery

We used a pre-trained convolutional model (without the softmax classification layer) as part of our model archiacea for bone age estimation. Analogously to the previous case, we fine-tuned OverFeat, GoogLeNet and OxfordNet on our X-ray image dataset.

2.3 BoNet: an ad-hoc CNN for skeletal bone age assessment

We trained a new CNN — called BoNet — trained on the X-ray scan dataset. The advantage of training a CNN over fine-tuning an existing one lies in the possibility to fine-tune the network architecture to the type of images at hand.

2.4 ***Experiment***

All models were trained using mini-batch stochastic gradient descent on an MSE loss function, with batch size set to 4, learn rate initially set to 0.002 and decay 0.0002. During validation and test, we provided a whole rescaled image as input to the CNN, obtaining as output a2D map of age estimations corresponding to different regions of the input image.
We built T = 8 trees with maximum depth D = 14,where for each node split 100 candidate features and 10 candidate thresholds were generated. In each round we randomly split the 60 input images into 43 training and 17 testing images. The measure we used for evaluating the performance is the Euclidean distance between the ground truth and the estimated landmark position.

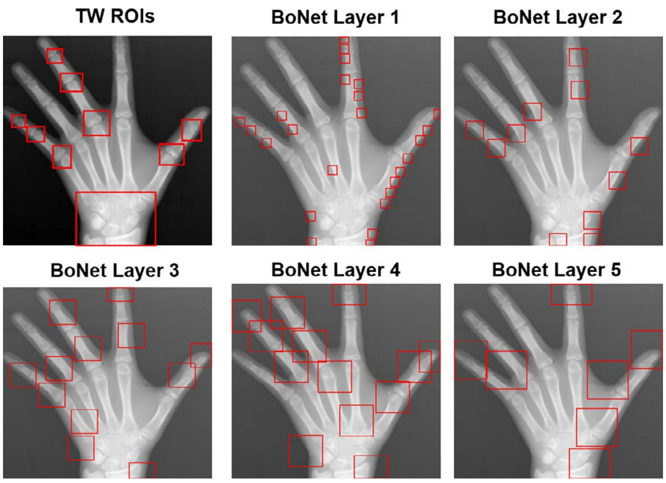

Figure 2: main localisation of CNN during Pre-Processing stage.

3. Results

CNN proved to be to accuracy for this project and the linear data estimation have been efficient as per the outputs.

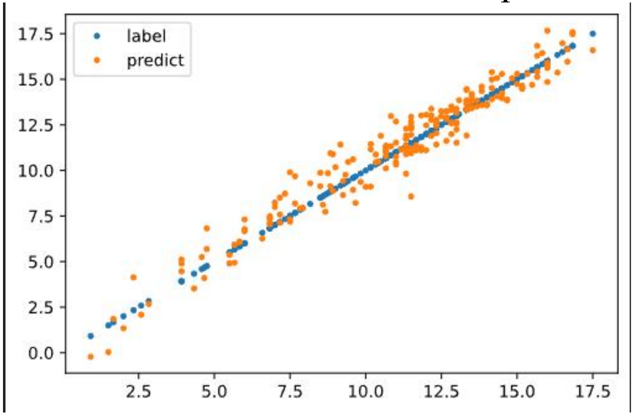

Figure 3: Training results.

Most of the existing methods, especially those based on the TW clinical methods, operate on local image patches (or regions) and do not need any pre-processing to extract some specific images. But training a CNN on specific images means identifying specific low- and middle-level features that, given the encouraging performance of BoNet, can be also useful for clinical procedures.

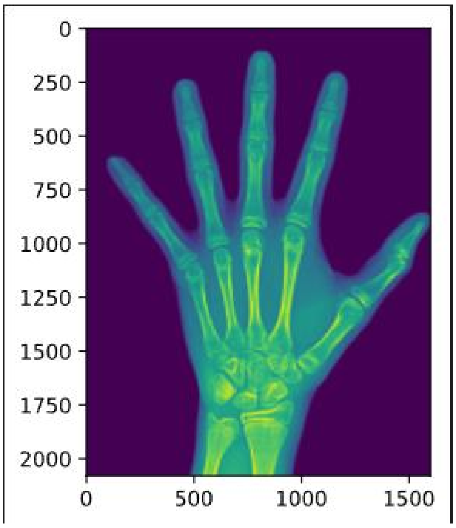

Figure 4: Hand image after CNN is applied.

4. Conclusion

The algorithm achieved the best mean error on radius and ulna bone, which can be explained by the large anatomical variation. In this experiment, there were at most three outliers in one single image, compared to others. On our dataset is able to achieve a much better accuracy when including the weighting function accordingly, compared to a weighting equal to one. CNN proved to be more efficient and effective when it comes to detection and estimation.

5. Reference List

Spampinato, C., Palazzo, S., Giordano, D., Aldinucci, M., & Leonardi, R. (2016, October 29). Deep Learning for Automated Skeletal Bone Age Assessment In X-ray Images. Deep learning for automated skeletal bone age assessment in X-ray images - ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S1361841516301840

Lee, H., Tajmir, S., Lee, J., Zissen, M., Yeshiwas, B. A., Alkasab, T. K., Choy, G., & Do, S. (2017, March 8). Fully Automated Deep Learning System for Bone Age Assessment - Journal Of Digital Imaging. SpringerLink. https://link.springer.com/article/10.1007/s10278-017-9955-8.

Wang, Y., Zhang, Q., Han, J., & Jia, Y. (2018, December 1). Application Of Deep Learning In Bone Age Assessment - IOPscience. Application of Deep learning in Bone age assessment - IOPscience. https://iopscience.iop.org/article/10.1088/1755-1315/199/3/032012.