Teeth category classification via seven-layer deep convolutional neural network with max pooling and global average pooling

Zhi Li | Shui-Hua Wang | Rui-Rui Fan | Gang Cao | Yu-Dong Zhang | Ting Guo

Prepared by Nesma Osama

Problem Explanation

This paper aims to classify teeth images into four categories: incisor, molar, canine, and premolar. Cone Beam Computed Tomography (CBCT) is used to produce high-quality images suitable for this classification task. While traditional machine learning techniques can be applied to solve this problem, they often require manual feature extraction. To address this limitation, this paper proposes using Convolutional Neural Networks (CNNs), which can automatically learn and extract features relevant to the classification task.

Dataset and Preprocessing:

They collected approximately 400 images 200 for training and 200 for testing. The dataset is balanced, meaning each class (incisor, molar, canine, and premolar) contains 50 images for training and 50 for testing. All images were converted to grayscale and resized to 64×64 pixels. To reduce the risk of overfitting, data augmentation techniques were applied to the training set. Each original training image was used to generate 180 augmented images.

Table 1. Setting of data augmentation

DA Approach	Parameters
Image rotation	Rotation angle θ = −15:1:15
Gamma correction	Gamma value r = 0.7:0.02:1.3
Noise injection	Variance = 0.01
Random translation	30 times
Scaling	Scaling factor s = 0.7:0.02:1.3
Random affine	30 times

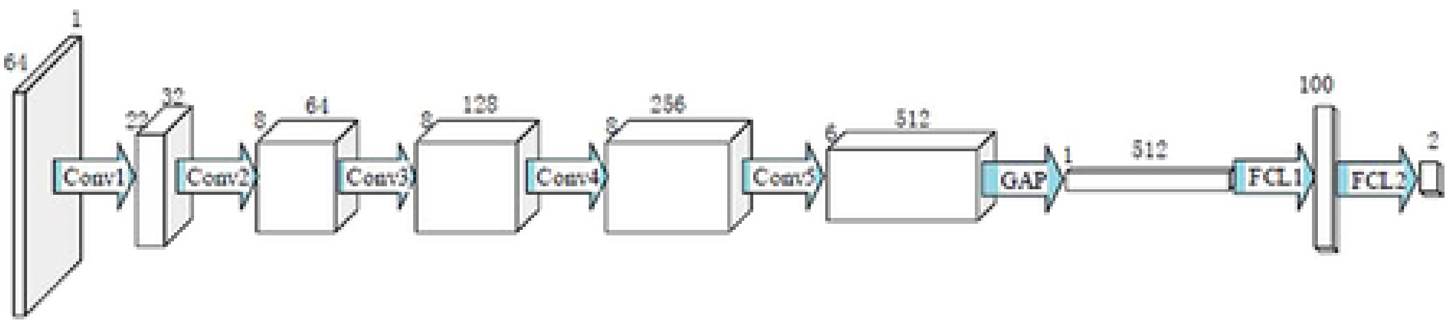
Methodology

The Convolutional Neural Network (CNN) used in this work is composed of alternating convolutional and pooling layers for feature extraction, followed by one or more fully connected layers for classification. Unlike traditional machine learning methods, which require manual feature design, CNNs can automatically learn and extract relevant features from the input data during training. This ability to learn features directly from images makes CNNs especially effective for image classification tasks like teeth type identification.

Neural network structure

The training data is fed into a seven-layer deep Convolutional Neural Network (CNN). The depth refers specifically to layers that contain learnable parameters ayers such as pooling, ReLU (activation), and softmax are not counted in this total, as they do not contain learnable weights or biases. Each convolutional layer is followed by a max-pooling (MP) layer to progressively reduce the spatial dimensions. After five convolutional layers, a global average pooling (AP) layer is applied to shrink the activation maps to 1×1. Finally, two fully connected layers with 100 and 2 neurons, respectively, are added for classification.

The structure of the 7-layer deep neural network



Conclusion

In the final experiment, the proposed seven-layer CNN was compared with several state-of-the-art methods on the same dataset. The result demonstrate that our CNN outperforms the other methods significantly in terms of classification accuracy.

Table 4. Comparison to state-of-the-art approaches in terms of sensitivities

Approach	C_1	C_2	C_3	C_4	Average
GOA (2017)8	76.0	78.0	76.0	84.0	78.5
LM (2016)9	78.0	82.0	74.0	84.0	79.5
HWT (2018)7	82.67	81.00	81.33	82.33	81.83
7-layer CNN (Our)	88.0	86.0	84.0	90.0	87.0