Title: Self-Supervised Clustering of Medical Images Using EfficientNet and KMeans

Name: Tara Setareh

Affiliation: Iran University of Science and Technology

Abstract: In this project, I explore a self-supervised approach to analyzing chest X-ray images using EfficientNet, a state-of-the-art convolutional neural network. Instead of traditional supervised classification, we utilize EfficientNet to extract deep image features and apply KMeans clustering to group the images without any label information. This method enables meaningful separation of medical images into categories (e.g., healthy vs. abnormal) even in the absence of annotated data. Our results suggest that EfficientNet-based feature extraction captures essential structural patterns that support effective clustering.

Introduction: Medical image classification typically relies on large sets of labeled data, which are costly and time-consuming to obtain. In scenarios where annotated data is scarce, self-supervised learning offers a promising alternative. This study proposes a simple yet effective self-supervised method for grouping medical images by combining EfficientNet feature extraction with KMeans clustering. The objective is to evaluate whether EfficientNet can be used in a non-supervised setting to separate X-ray images into semantically meaningful groups.

Methodology: We used the pretrained EfficientNet-B0 model to extract 320-dimensional feature vectors from chest X-ray images. Images were first preprocessed using standard transformations (resizing, normalization). We then applied adaptive average pooling on the final convolutional output to obtain a fixed-size feature vector for each image. These vectors were input to the KMeans clustering algorithm (k=2) to group the images. The experiment was implemented in Python using PyTorch and Scikit-learn libraries. A dataset of 20 X-ray images, including normal and abnormal samples, was used for testing.

Results: The clustering process successfully grouped the images into two distinct clusters. Visual inspection of the images in each cluster revealed that similar types of images (e.g., normal chest vs. infected) tended to be grouped together. This indicates that EfficientNet features, though not trained for this specific task, can still capture latent semantic similarities. The method demonstrated the potential for initial exploration and organization of unlabeled medical image datasets.

Conclusion: This study presents a lightweight and efficient self-supervised approach to clustering medical images using EfficientNet and KMeans. Even without labels, the model was able to learn meaningful representations that allowed clustering based on visual similarity. This method could be used as a preliminary step in medical image analysis workflows, especially where labeling is expensive or unavailable.

Keywords: Self-Supervised Learning, EfficientNet, Medical Imaging, Feature Extraction, KMeans Clustering, Chest X-ray

```python
import torch
from torchvision import transforms
from efficientnet_pytorch import EfficientNet
from PIL import Image
import os
import numpy as np
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt
```

[3] ✓ 3.2s Python

```python
model = EfficientNet.from_pretrained('efficientnet-b0')
model.eval()
```

[4] ✓ 33.4s Python

Figure 1:Initialization of the EfficientNet-B0 model, loaded with pretrained weights for feature extraction on medical images.

```python
for i, img_name in enumerate(os.listdir(image_dir)):
    print(f"{img_name} → cluster: {labels[i]}")
```

[11] ✓ 0.0s Python

```
189.jpg → cluster: 0
374.PNG → cluster: 1
818-.png → cluster: 1
162.jpg → cluster: 0
252.jfif → cluster: 1
638.png → cluster: 0
610.png → cluster: 0
604.png → cluster: 0
837.png → cluster: 1
348.jpg → cluster: 0
360.jpg → cluster: 1
412.jpg → cluster: 0
770.png → cluster: 0
228.jpg → cluster: 0
200.jpg → cluster: 1
566.png → cluster: 0
214.png → cluster: 0
572.png → cluster: 1
599.png → cluster: 0
109.jpeg → cluster: 1
482p-.jpg → cluster: 1
244.jfif → cluster: 0
765.jpg → cluster: 1
215.png → cluster: 0
573.png → cluster: 0
```

Figure 2:Final clustering output, listing chest X-ray images and their corresponding cluster assignments (Cluster 0 or 1). This unsupervised grouping is based purely on features extracted by EfficientNet without using any labe.

Figure 3: Sample chest X-rays from different clusters. The top image (Cluster 1) highlights a potential abnormality with an arrow marker, while the bottom image (Cluster 0) appears relatively normal. This visual comparison supports that the model successfully grouped similar patterns without any labels.

Source of EfficientNet code: https://github.com/lukemelas/EfficientNet-PyTorch