

Case ID:M20-069L^

Published: 8/18/2020

Inventors

Mohammad Reza Hosseinzadeh

Taher

Fatemeh Haghighi

Jianming Liang

Contact

Jovan Heusser
jovan.heusser@skysonginnovations.com

Transferable Visual Words: Chest X-ray Image Analysis

Image analysis techniques are invaluable in helping physicians better diagnose and treat diseases and expand the utility of medical imaging. Two common image analysis techniques, convolutional neural networks (CNN) and bags of visual words (BoVW) are often thought of as competing methods, however, they could have complementary strengths. CNNs are great at object recognition and classifications, but they require a significant amount of training images. The learned features, however, are transferrable to other applications. BoVW treats image features as words, is unsupervised and does not require expert annotation, which is tedious and expensive, but it cannot be fine-tuned.

Researchers at Arizona State University have developed a novel method, called Transferable Visual Words (TransVW) for chest X-ray image analysis. This method integrates CNNs and BoVW to amplify their strengths and overcome some of their limitations. TransVW uses the transfer learning capability of CNNs with the unsupervised nature of BoVW in extracting visual words, and results in a new self-supervised method. When TransVW was evaluated on an NIH hospital-scale chest X-ray dataset, it outperformed all of the state-of-the-art approaches, including fine-tuning pre-trained ImageNet models, which is a significant accomplishment.

This annotation efficient method combines the power of transfer learning with CNNs and the unsupervised property of visual word extraction with BoVW leading to a new and superior self-supervised method, TransVW, for image analyses.

Potential Applications

- Chest X-ray image analyses
 - o Atelectasis, cardiomegaly, effusion, infiltration, nodule, pneumonia, pneumothorax, consolidation, edema, emphysema, fibrosis, pleural thickening, hernia, and more

Benefits and Advantages

- Self-supervised, requires no expert annotation in pre-training

- Conceived from a new idea that the sophisticated, recurrent anatomical structures in medical images are natural visual words
- The natural visual words are automatically extracted and serve as strong, but supervision-free signals for CNNs to learn generalizable image representation
- Automatically extracts visual words directly from X-ray images, ensuring their consistency in appearance
- New U-Net-like architecture to enhance representation learning capability
- Outperforms all of the state-of-the-art approaches including fine-tuning pre-trained ImageNet models
- Reduces annotation efforts by 75% relative to training model from scratch and by 12% relative to fine tuning a pre-trained ImageNet model
- Significantly accelerates the convergence speed in comparison with training from scratch and ImageNet-based transfer learning

For more information about the inventor(s) and their research, please see

Dr. Liang's departmental webpage