Vision Vault

Team Members:
Abhianash Gunda(RA2111003010850)
Hemanth Ettamsetti(RA2111003010480)
Taruni Mamidipaka(RA2111003010888)

Abstract

Image classification is a commonly used biometric technique for item and face recognition, with face recognition making extensive use of its natural and non-intrusive approach. Deep learning networks with Triplet Loss have gained popularity recently as a standard framework for identity verification and human identification. We describe a novel approach in this study for choosing suitable hard-negatives for Triplet Loss training. It is demonstrated that adding pairs that would have been eliminated otherwise improves accuracy and performance. Moreover, we used the Adaptive Moment Estimation approach to reduce the possibility of early convergence brought on by the extra hard-negative pairs. In open face, we were able to obtain an accuracy of 0.968, while in LBPH, we saw substantially lower accuracy.

Introduction

Image Classification (IC) finds use in multiple fields. It highlights how important it is to accurately label and classify images using well-trained algorithms that produce few errors. Highlighted are the applications of IC in various industries and daily life, demonstrating its effectiveness in a range of fields.

It is known that Image Processing (IP) and Computer Vision (CV) techniques play a significant role, especially in face identification. The computing requirements of IC datasets in comparison to text-based data are also mentioned in the text.

A brief explanation is given of two popular IC approaches: single-class (SC) and multi-class (UC) learning. Three steps make up the Model Structure process: creating the model architecture, training the learning model, and stacking and preprocessing the dataset.

It emphasizes that well-organized datasets are essential to reliable model results. Selecting the quantity of layers and hidden units is a step in the process of creating a model architecture. Labeled figures are placed for both training and validation as part of the learning model training process.

Loading test data and applying the trained model to class prediction constitute the evaluation phase. All in all, the article offers a succinct summary of the essential elements of image classification, including its applications and the procedures involved in creating and assessing a successful model.

Problem Statement

The text addresses challenges in Computer Vision (CV), particularly in Image Classification (IC) and Object Detection (OD). With the rise of Deep Learning (DL) methods, precision has improved. The focus is on IC, involving creating a labeled dataset for image localization and bounding box formation around detected objects, enhancing accuracy and efficacy. Object Detection presents critical challenges, including both classification and localization. Facial feature extraction is crucial for identifying objects, particularly in recognizing individuals from face images. The model extracts faces, corrects features, and stores them for further validation, contributing to automatic machine verification. The dataset contains features for improved images obtained during registration, ensuring consistency during validation.

EXISTING MODEL AND PROPOSED MODEL

Existing Model: Current image classification models often use traditional machine learning algorithms or early convolutional neural networks (CNNs). However, they may struggle with complex patterns, diverse datasets, and achieving high accuracy, especially relying on handcrafted features.

Proposed Model: The proposed image classification model adopts advanced deep learning techniques, including state-of-the-art CNN architectures or newer models like Transformers. These models automatically learn hierarchical features, improving accuracy on diverse datasets. The proposed model may also employ transfer learning, leveraging pre-trained models for enhanced performance on specific tasks with limited data.

MOTIVATION WITH SOCIETAL BENEFIT

Advancements in image classification offer substantial societal benefits. From early disease detection in medical imaging to improved security and surveillance, environmental monitoring, safer autonomous vehicles, enhanced agriculture, accessibility tools for the visually impaired, interactive education, and wildlife conservation, the technology positively impacts various domains, leading to better health, safety, sustainability, and inclusivity.

ARCHITECTURE OF PROPOSED MODEL

The proposed image classification model employs a sophisticated deep learning architecture, likely a Convolutional Neural Network (CNN) or Transformer-based model. Key components include convolutional and pooling layers for feature extraction, normalization, activation functions, attention mechanisms (if applicable), fully connected layers, and optional regularization techniques. Transfer learning from pre-trained models enhances adaptability and scalability. The architecture is designed for superior accuracy and efficiency in handling diverse image classification tasks.

MOTIVATION WITH SOCIETAL BENEFIT

1. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks.
2. Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition.
3. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition.
4. Russakovsky, O., et al. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV), 115(3), 211-252.
5.Huang, G., Liu, Z., & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks.

THANK YOU