Int J Syst Assur Eng Manag
https://doi.org/10.1007/s13198-024-02345-4

ORIGINAL ARTICLE

From scratch or pretrained? An in-depth analysis of deep learning approaches with limited data

Saqib Ul Sabha[1] · Assif Assad[1] · Nusrat Mohi Ud Din[1] · Muzafar Rasool Bhat[2]

Received: 18 November 2023 / Revised: 28 March 2024 / Accepted: 7 April 2024

Abstract The widespread adoption of Convolutional Neural Networks (CNNs) in image recognition has undeniably marked a significant breakthrough. However, these networks need a lot of data to learn well, which can be challenging. This can make models prone to overfitting, where they perform well on training data but not on new data. Various strategies have emerged to address this issue, including reasonably selecting an appropriate network architecture. This study delves into mitigating data scarcity by undertaking a comparative analysis of two distinct methods: utilizing compact CNN architectures and applying transfer learning with pre-trained models. Our investigation extends across three disparate datasets, each hailing from a different domain. Remarkably, our findings unveil nuances in performance. The study reveals that using a complex pre-trained model like ResNet50 yields better results for the flower and Maize disease identification datasets, emphasizing the advantages of leveraging prior knowledge for specific data types. Conversely, starting from a simpler CNN architecture trained from scratch is the superior strategy with the Pneumonia dataset, highlighting the need to adapt the approach based on the specific dataset and domain.

Keywords Pneumonia detection · Limited data · CNN · Transfer learning · Pre-trained model · Flower recognition · Maize disease identification

1 Introduction

DEEP learning has gained widespread recognition for its outstanding performance and adaptability across various domains. These domains encompass a multitude of applications, including Object Recognition (Hu et al. 2018), image classification (Jin et al. 2022; Krizhevsky et al. 2017; Dar et al. 2022), object detection (Ren et al. 2015; Dar et al. 2023), image segmentation (Long et al. 2015) and many other fields. In the context of image classification tasks, the objective is to assign each image to a specific category. This problem involves taking an image as input and assigning a corresponding label. With advancements in high-performance computing, more advanced Convolutional Neural Networks (CNNs) have been introduced for image classification, examples being Alexnet (Krizhevsky et al. 2012) and VGG (Simonyan and Zisserman 2014). Typically, CNN architectures comprise convolutional and pooling layers, with both operations enhancing the robustness of feature representations. Furthermore, some CNNs are designed to handle additional supervised learning tasks. For instance, a fully convolutional network introduced in Long et al. (2015) replaces fully connected layers in the top layers of traditional CNNs with fully convolutional layers. The CNN is a sophisticated architecture for learning representations across multiple layers. It has achieved remarkable success in various domains,

✉ Saqib Ul Sabha
saqib.sabha@islamicuniversity.edu.in

Assif Assad
assifassad@gmail.com

Nusrat Mohi Ud Din
nusratmohiuddin92@gmail.com

Muzafar Rasool Bhat
muzafar.rasool@islamicuniversity.edu.in

1 Department of Computer Science and Engineering, Islamic University of Science and Technology, Pulwama, Jammu and Kashmir 192122, India

2 Department of Computer Science, Islamic University of Science and Technology, Pulwama, Jammu and Kashmir 192122, India