Testing Transfer Learning for Image Clasification

Abhijith Dameruppala

ABSTRACT

This project investigates the efficiency of transfer learning with datasets in the context of image classification. Utilizing a custom dataset, we investigate the minimal number of image samples required to achieve a predetermined level of accuracy across various pre-trained models: VGG-16, ResNet-50, MobileNet, and EfficientNet. This research not only addresses the practical constraints of data collection in machine learning projects but also explores the adaptability and efficiency of transfer learning across different models. The core methodology of this project involves the careful curation of a unique dataset comprising images of smartphones and remotes, characterized by their diversity in design, orientation, and context. Through this investigation, the project seeks to advance the understanding of transfer learning's limits and possibilities, providing a foundational framework for future research in the field.

***Index Terms*— Data Augmentation, Dataset Curation, Deep Learning Models, EfficientNet, Feature Extraction, Image Classification, Model Fine-Tuning, MobileNet, Remote Controls, ResNet-50, Smartphones, Transfer Learning, VGG-16.**

I. INTRODUCTION

Imagine trying to explain to a friend over the phone which remote on the coffee table turns on the TV, and which one is your phone lying face down. It sounds simple, but visually, these objects share a lot of similarities—flat, rectangular, and filled with buttons. This everyday confusion sets the stage for our project, diving into the nuanced world of image classification where we tackle distinguishing between smartphones and remote controls using the magic of machine learning.

At the heart of our exploration is a technique known as transfer learning. It's a bit like teaching an experienced chess player how to play checkers. They don't start from scratch; instead, they adapt their rich understanding of board games to master a new game more quickly. Similarly, transfer learning allows us to take a model that's already great at recognizing objects and fine-tune it to become an expert in spotting the differences between our two chosen objects.

But why smartphones and remotes? Apart from them often being mistaken for one another, they represent a broader challenge in machine learning: making accurate calls when data is scarce. Whether it's in healthcare, where each data point is precious, or in environmental science, where rare species are hard to catch on camera, the need to learn a lot from a little is everywhere.

So, we're building our own collection of smartphone and remote images. This dataset will be used for testing out some of the smartest models such as VGG-16, ResNet-50, MobileNet, and EfficientNet, to see just how little they need to learn a lot. Through this project, we aim to shed light on the efficiency of transfer learning and how it might be the best in tackling data scarcity across various fields.

II. METHODOLOGY

DATASET CREATION AND PREPROCESSING:

We plan to gather a diverse set of images for both smartphones and remote controls. This will include images from online repositories, open-source datasets, or photographs taken specifically for this project. Preprocessing the dataset will include standardizing the image sizes, and performing data augmentation (rotation, flipping, scaling) to increase the robustness of the model.

SELECTON OF PRETRAINED MODELS:

Based on the literature review, and the nature of our project, we are considering the following models for their effectiveness in image classification tasks:

1. VGG16 or VGG19
2. ResNet-50
3. InceptionV3
4. MobileNet

TRANSFER LEARNING STRATEGY:

- Use the convolutional base of the pre-trained models to extract features from our dataset, keeping the model's pre-trained weights frozen. Add new classification layers tailored to our binary classification task (smartphone vs. remote control).
- After initial training with the feature extraction approach, selectively unfreeze some of the top layers of the convolutional base and jointly train both the newly added layers and these top layers to fine-tune the model on your specific dataset.

MODEL TRAINING:

Split our dataset into training, validation, and test sets to evaluate the model's performance. Compile the model, selecting an appropriate optimizer (e.g., Adam) and loss function (e.g., binary crossentropy for binary classification).

EVALUATION AND TESTING:

Evaluate the model's performance on the validation set during training to fine-tune hyperparameters and training configurations. And upon satisfactory validation performance, assess the final model on the test set to report metrics such as accuracy, precision, recall, and F1 score.

MODEL OPTIMIZATION:

Based on the initial results, iteratively adjust the model architecture, training process, or transfer learning strategy to improve performance. This may involve experimenting with different models, adjusting the number of layers unfrozen during fine-tuning, or varying the data augmentation techniques used.

III. LITERATURE REVIEW

Transfer Learning for Image Classification Using VGG19: Caltech-101 Image Dataset

Objective: The primary objective of the study was to improve image classification performance by merging deep features extracted using the VGG19 convolutional neural network with traditional handcrafted feature extraction methods, including SIFT, SURF, ORB, and the Shi-Tomasi corner detector algorithm. The study hypothesized that a hybrid feature extraction approach could overcome limitations observed solely with deep learning-based or traditional methods, thereby enhancing classification accuracy.

Methodology: The authors employed a structured approach encompassing feature extraction, dimensionality reduction, and classification phases. The VGG19 model was utilized to extract deep features, while SIFT, SURF, ORB, and Shi-Tomasi algorithms were applied for handcrafted feature extraction. To manage the high-dimensional feature space, techniques like k-means clustering and Locality Preserving Projection (LPP) were used for dimensionality reduction. The resulting features were then classified using several machine learning algorithms, including Gaussian Naïve Bayes, Decision Tree, Random Forest, and XGB Classifier. The experimentation was conducted on the Caltech-101 dataset, known for its complexity and variability across image classes.

Findings: The study revealed that the Random Forest classifier, when used with the combined features, achieved the highest accuracy of 93.73%, outperforming other classifiers and methods proposed by different researchers. This finding underscores the effectiveness of integrating deep learning features with handcrafted features for image classification tasks.

Project Implications: The findings and methodology of Bansal et al. inform our project on distinguishing between smartphones and remotes. Particularly, their approach to combining deep learning with traditional feature extraction techniques could enhance our model's accuracy, especially given the visual similarities between the objects of interest. Their successful use of VGG19 alongside traditional methods like SIFT and SURF suggests a promising direction for our project's methodology, potentially enabling our model to capture a more comprehensive feature set from the images.

Transfer Learning for Medical Image Classification: A Literature Review

Objective: The objectives of this paper are to identify key methodologies, successes, limitations, and areas requiring further investigation within this domain. And to offer insights into future directions for research and practical applications in medical imaging analysis.

Methodology: Kim et al. systematically reviewed 425 peer-reviewed articles from PubMed and Web of Science up to December 31, 2020, following PRISMA guidelines. A total of 121 studies met their inclusion criteria, focusing on the selection of backbone models and TL approaches including feature extraction, feature extraction hybrid, fine-tuning, and fine-tuning from scratch.

Findings:

- A significant number of studies (57) empirically evaluated multiple models, with Inception being the most employed. Deep models, particularly ResNet and Inception, were recommended for their efficiency as feature extractors.
- Multiple approaches were empirically benchmarked in most studies (46), with feature extractor and fine-tuning from scratch being the most favored approaches due to their balance between performance and computational cost.
- The review observed a wide variety of data characteristics across the studies, including image modality, data subject, and dataset size, which influenced the choice of TL approach.

Project Implications: For our project on distinguishing between smartphones and remote controls using transfer learning, the insights from Kim et al.’s review could be invaluable:

- Consider deep learning models like ResNet or Inception for their demonstrated effectiveness in medical image classification, which may translate well to our task.
- Begin with feature extraction to leverage pre-trained models efficiently and explore fine-tuning for further optimization based on your dataset’s specific characteristics.
- Reflect on your dataset's uniqueness in terms of size, variety, and complexity when adapting the TL strategies highlighted in the review.

Deep Transfer Learning for Image Classification: A Survey

Objective: The survey extensively reviews the application of deep transfer learning to image classification, addressing the challenge of achieving high accuracy in image classification tasks when there are constraints on the availability of large, labeled datasets.

Methodology: The survey clarifies the process of transfer learning, specifically deep transfer learning, where knowledge gained from training on a large source dataset (DS) is applied to a new, but related, target task (TT) to improve performance despite data limitations. Over 425 peer-reviewed articles were reviewed, focusing on the selection of backbone models, transfer learning approaches (feature extraction, fine-tuning), and their applicability across diverse image classification scenarios.

Findings: The survey highlights the predominant use of deep models like Inception and ResNet as effective feature extractors in transfer learning, noting their efficiency in handling image classification tasks. Insights into how the nature of the dataset (image modality, subject matter, and size) influences the choice of transfer learning strategy and model architecture. The paper reports on the effectiveness of feature extraction and fine-tuning from scratch as favored approaches, balancing performance with computational efficiency.

Project Implications:

- Consider deep learning models known for their robust feature extraction capabilities, as identified in the survey.
- The insights on data characteristics can guide the development of your dataset, ensuring it's well-suited for transfer learning applications.
- Leverage the survey's findings on the effectiveness of various transfer learning approaches to optimize your methodology, potentially exploring both feature extraction and fine-tuning strategies based on your dataset size and similarity to the source dataset.

REFERENCES

[1] Kim, H.E., Cosa-Linan, A., Santhanam, N. et al. Transfer learning for medical image classification: a literature review. BMC Med Imaging 22, 69 (2022). https://doi.org/10.1186/s12880-022-00793-7

[2] Bansal, M., Kumar, M., Sachdeva, M. et al. Transfer learning for image classification using VGG19: Caltech-101 image data set. J Ambient Intell Human Comput 14, 3609–3620 (2023). https://doi.org/10.1007/s12652-021-03488-z

[3] Plested, Jo & Gedeon, Tom. (2022). Deep transfer learning for image classification: a survey. 10.48550/arXiv.2205.09904.

[4] GfG. (2023, January 10). Residual networks (resnet) - deep learning. GeeksforGeeks. https://www.geeksforgeeks.org/residual-networks-resnet-deep-learning/

[5] Awan, A. A. (2022b, November 23). A complete guide to data augmentation. DataCamp. https://www.datacamp.com/tutorial/complete-guide-data-augmentation

[6] “Dataset Creation and Curation.” Introduction to Data-Centric AI, dcai.csail.mit.edu/2024/dataset-creation-curation/. Accessed 1 Apr. 2024.

[7] Biswal, Avijeet. “Top 10 Deep Learning Algorithms You Should Know in (2021).” Simplilearn.com, 16 Feb. 2023, www.simplilearn.com/tutorials/deep-learning-tutorial/deep-learning-algorithm.

[8] Tan, Mingxing and Quoc V. Le. “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.” ArXiv abs/1905.11946 (2019): n. pag.

[9] Agarwal, Dishaa. “Feature Extraction | Popular Feature Extraction Techniques.” Analytics Vidhya, 19 Apr. 2021, www.analyticsvidhya.com/blog/2021/04/guide-for-feature-extraction-techniques/.