International Journal of Computer Communication and Informatics

DOI: 10.34256/ijcci2325

Harnessing Effectiveness of ResNet-50 and EfficientNet for Few-Shot Learning

Santoshi Vajrangi [a,*], Satvikraj Selar [a], Anupama P Bidargaddi [a], Rishi Hiremath [a], Vivek Yeli [a]

[a] School of Computer Science and Engineering, KLE Technological University, Karnataka, India

* Corresponding Author Email: santoshivajrangi@gmail.com

Received: 29-10-2023, Revised: 03-12-2023, Accepted: 15-12-2023, Published: 24-12-2023

Abstract: Inspired by the concept of human intelligence- learning and expanded upon with several examples – several- step learning focused on computers that can classify images in a comparable way. This article covers the interesting field of sparse learning, focusing on comparing its implementation using two popular deep learning networks: ResNet and EfficientNet. Little learning has the potential to be effective on tasks where obtaining large data sets is expensive or impossible. This allows machines to mimic humans' ability to learn and expand from small samples, thus opening possibilities in the field of several types of diagnostics, personalized recommendations, systems, and robotics. Our main goal is to measure and compare the accuracy achieved by these models when learning on limited datasets and to show that EfficientNet achieves better accuracy when it requires fewer parameters and computational resources compared to ResNet-We considered VGG-flowers dataset for comparison. Our results show that Narrow EfficientNet outperforms ResNet-50 in terms of overall accuracy (85.20% vs. 84.30%), precision (85.60% vs. 85.40%), recall (85.30% vs. 84.50%) and F1 acquisition (85.45% vs. 84.95%). This suggests that EfficientNet's emphasis on computational efficiency and parallelism may provide a slight advantage on limited data.

Keywords: Few-shot learning, ResNet-50, EfficientNet, Image classification

1. Introduction

ResNet and EfficientNet represent two pillars of convo- lutional neural network (CNN) design, each with unique advantages in processing complex visual data. This study aims to evaluate the accuracy and efficiency of ResNet and Effi- cientNet [1] on various classification systems. Convolutional neural networks (CNN) have played an important role in the advancement of computer vision, especially in tasks such as image classification. Among the many CNN architectures, ResNet and EfficientNet have become two important sources of data visualization, each good for data visualization.

ResNet is an abbreviation for Residual Network, which combines residuals generated using CNN. The blocks are equipped with short joints to create good slopes from depth.

This innovation solves the problem of vanishing gradients and enables deep model training. ResNet has become a standard for demonstrating the effectiveness of image classification models in deep learning.

EfficientNet focuses on the efficient use of IT resources. It uses different methods to adjust the depth, width, and resolution of gridcitb2. This approach shows that the model with performance measurements is not better than the original model. EfficientNet achieves significant advantages in terms of rendering technology and computational efficiency, making it an attractive choice for visualization applications.

The importance of ResNet and EfficientNet extends beyond the traditional image classification scenario. In this work, we examine various aspects of the scientific model, where simulating real situations is a challenging task, and the model must perform better than other models. Multi-learning is important in applications where acquiring knowledge in more than one domain is impractical or costly [2-4].

The main purpose of this study is to evaluate and compare the accuracy and efficiency of ResNet and EfficientNet in different scenarios. Deployment scenarios. Our aim is to understand the evolution of models in insufficient informa- tion situations by evaluating their performance in insufficient information situations.

Through comparison, our goal is to understand how ResNet and EfficientNet can be modified to meet the needs of research as a representative CNN. The results of this study have implications for many fields, from medical image analysis to eye examinations, and it is possible to learn from important examples.

After rigorous testing and careful analysis, we found that we have a good understanding of the performance of ResNet-50 and EfficientNet in multi-hour deployment scenarios [5]. Our results show that EfficientNet slightly outperforms ResNet-50 in terms of overall accuracy (85.20% vs. 84.30%) and accuracy (85.60% vs. 84.30%), recovery (85.30% vs. 84.50%) and F1 gain (85.45% vs. 84.95%). This suggests that EfficientNet's emphasis on computational efficiency and parallelism may provide a slight advantage on limited data.

2. Background Study

Research on the distribution of small samples has many needs. Research shows that the workplace is making signifi- cant progress with improved accuracy, data availability, and new training or models. First, the processing of data is demonstrated on a Dual Channel Prototype Network (DCPN) using Transformers and CNN for classification of various diseases, which represents a significant improvement over the basic principles(previousb6).

Secondly, a new way to use multiple views is proposed which is "Cross", a few-model hyperspectral correlation learning for image classification. This method uses multiple learning bases with random learning using future extractors and interactive models. It provides different learning methods at the category level, cross-tracking at the setting level for classification at various levels, and isolation at the domain level. Development of hyperspectral image classification. This approach emphasizes cross-matching, resulting in better performance [7, 8] in tests. Also, by solving new categories. It uses Proto-type Former (a Transformer-based approach) to discover social patterns, improve features through random learning, and attempt to surpass state-of-the-art drawing methods to be more efficient and accurate [9].

Finally Few SAR takes a good approach. Few SAR aims to solve the lack of integration of SAR image classification by providing an easily accessible benchmark consisting of 15 methods. Goals include Road Construction, road analysis and planting.

2.1 Model Architectures

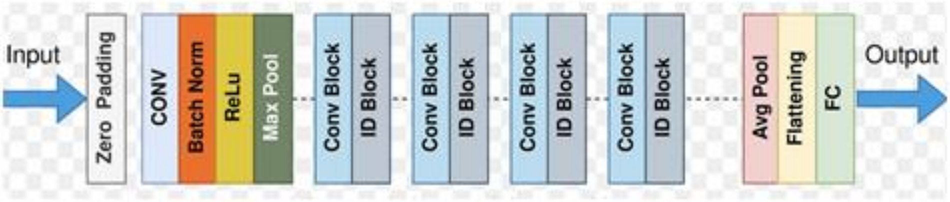

Figure 1. ResNet-50 Architecture

1) *ResNet-50 Architecture:* ResNet-50, short for 50-layer residual network, is a deep convolutional neural network architecture designed to solve the problem of vanishing in deep neural networks. The main innovation of ResNet is the introduction of fast-connecting components that allow the model to traverse one or more layers. This allows training of very deep networks without any degradation issues.

Architecture consists of several parts, each consisting of two or more layers which helps the model understand the business, making it easier to build and train more in-depth. Cross-matching allows messages to flow directly from one layer to another, thus supporting matching of gradients during backpropagation.

ResNet-50 now has a deep model that is content-dependent, proven and standard for image classification.

At the heart of ResNet-50 is a unique design: connectivity and "cross-connections" between layers. These clever shortcuts bypass layers and allow important messages to flow throughout the network. This design process not only enables ResNet-50 to achieve deep learning, but also facilitates effec- tive training even with such models [10]. As a result, it achieves good accuracy, ensuring reliable image distribution.

But the power of ResNet-50 comes at a price. Its complex architecture requires huge computing power to operate [11]. The number of layers means a large number of parameters, pushing the limits of computational feasibility on some platforms [12]. Therefore, although ResNet-50 is superior in terms of accuracy and is a reliable benchmark for comparison, its capabilities should be carefully evaluated before deployment [13].

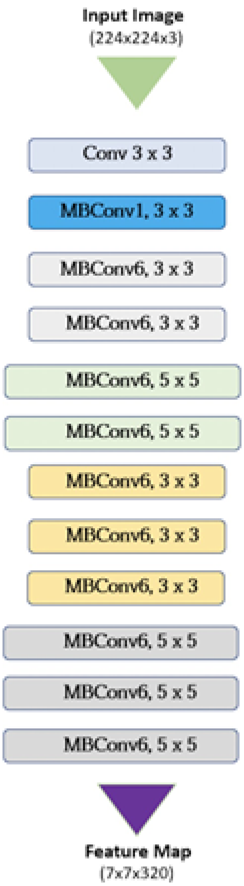

Figure 2. EfficientNet Architecture

2) EfficientNet Architecture: EfficientNet is a system designed to achieve high performance standards while using efficient computing resources. It provides a way to measure network depth, width, and resolution simultaneously. This approach allows EfficientNet to achieve high accuracy with fewer errors than other models.

EfficientNet has many blocks, and they all have the same backend as MobileNetV2. This model is parameterized by three parameters: α (width), β (depth) and γ (resolution). These measurements determine the sample size, allowing EfficientNet to measure variance.

EfficientNet provides a good balance between model ac- curacy and computational efficiency, making it suitable for a wide range of applications and deployments. Its multitasking performance makes it a popular choice for the computer vision impaired.

In summary, EfficientNet achieves its effectiveness by measuring the length of the network and providing different types of feature classification for images.

However, EfficientNet's early age is a cause for concern. Unlike the well-known ResNet-50, noise protection is closely examined. But EfficientNet's combination of accuracy and efficiency makes it a rising star and offers exciting potential for future image classification work.

3. Methodology

The methodology includes experimentation with a variety of few-shot learning scenarios, ranging from one-shot to a few- shots, to mimic real-world situations with limited labeled data[3]. A thorough measurement and analysis of the accuracy attained by each of the designs in various scenarios.

3.1 Dataset Description

1) VGG-Flowers Dataset: This dataset comprises 4242 images of flowers collected from various sources, including Flickr, Google Images, and Yandex Images for plant recognition tasks based on photographs. As shown in Figure 3, the dataset is classified into five classes namely Chamomile, Tulip, Rose, Sunflower, Dandelion, each representing a different type of flower, and each class has 800 images.

Figure 3. VGG- flowers Dataset (images)

2) Image Characteristics:

- Resolution: The images are not high resolution, with dimensions around 320x240 pixels.
- Size Variation: Photos are in their original size and not converted to a single size, and they exhibit different proportions.

Researchers and professionals can use it to study floral knowledge, and its size, resources, and characteristics make it suitable for many applications.

3.2 ResNet-50 Model Formulation

Feature Extraction: Let $f_{ResNet}(x)$ denote the feature extraction function of ResNet-50. For an input image x, the features are extracted as follows:

$$F_{ResNet} = f_{ResNet}(x) \tag{1}$$

ResNet-50 employs a series of convolutional layers to extract meaningful features [13] from the input image x. These layers act as filters, identifying patterns like edges, shapes, and textures that are crucial for visual understanding. The extracted features are encoded in the feature vector F, which represents a condensed, informative representation of the image content.

Classification Layer: The ResNet-50 features are then fed into a classification layer for the final prediction. Let W_cResNet and b_cResNet be the weight matrix and bias vector of the classification layer, respectively. The final prediction y_{ResNet} is obtained as:

$$y_{\ ResNet} = softmax(W_{cResNet} \cdot F_{ResNet} + b_{\ cResNet}) \tag{2}$$

The softmax function ensures the output probabilities sum to 1, representing a valid probability distribution over the classes. The final prediction y_{ResNet} indicates the model's confidence in each class.

3) *Loss Function:* The ResNet-50 model is trained using cross-entropy loss, which is evaluated using equation 3 for a single instance as given by:

$$L_{ResNet} = -\sum_{i=1}^{N} ti\ log\left(y_{ResNet,i}\right), \tag{3}$$

where N is the number of classes, t_i is the true label (in one- hot encoding), and $y_{ResNet,i}$ is the predicted probability of class i.

Cross-entropy loss measures the discrepancy between the model's predictions $y_{ResNet,i}$ and the true labels t_i. It penalizes incorrect predictions more harshly, guiding the model towards learning accurate class probabilities. The goal during training is to minimize this loss function, leading to better model performance.

3.3 EfficientNet Model Formulation

1) Feature Extraction: Let fEffNet(x) denote the feature extraction function of EfficientNet. The features of an input image x are extracted as follows: EfficientNet employs a carefully designed architecture that balances depth, width, and resolution to achieve high accuracy with fewer parameters.

$$F_{Eff\,Net} = f_{Eff\,Net}(x) \tag{4}$$

The feature extraction process [14] follows a similar concept as ResNet-50, using convolutional layers to extract informative features from the input image *x*. However, EfficientNet's architecture is specifically optimized for efficiency and com- pactness.

2) *Classification Layer:* The EfficientNet features are then fed into a classification layer for the final prediction. Let W_cEffNet and b_cEffNet be the weight matrix and bias vector of the classification layer, respectively. The final prediction y_{EffNet} is obtained as:

$$y_{EffNet} = softmax\left(W_{cEffNet} \cdot F_{EffNet} + b_{cEffNet}\right) \tag{5}$$

3) Loss Function: EfficientNet uses cross-entropy loss to measure the difference between predictions and true labels, guiding model training. The EfficientNet model is trained using cross-entropy loss, which for a single instance is given by:

$$L_{EffNet} = -\sum_{i=1}^{N} ti \; log\left(y_{ResNet,i}\right) \tag{6}$$

where N is the number of classes, t_i is the true label (in one- hot encoding), and $y_{EffNet,i}$ is the predicted probability of class i.

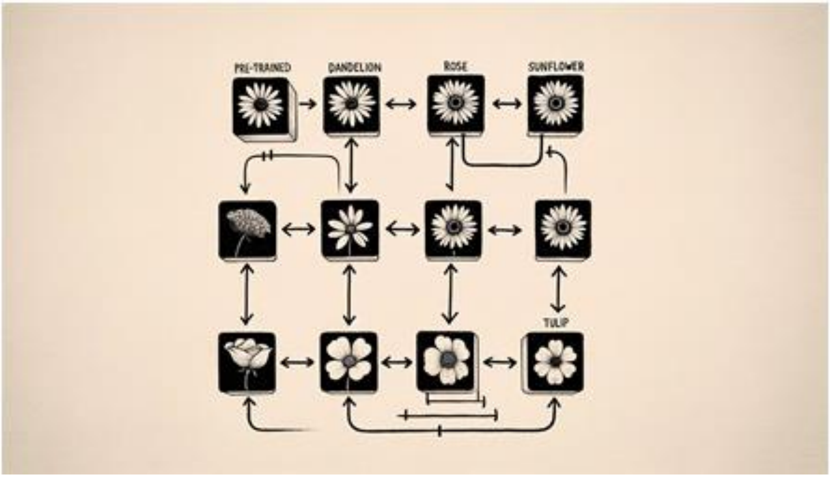

Figure 4. Few shot classification

4. Results Analysis

4.1 Performance Comparison

The table 1 below compares the performance of different models.

Table 1. Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score
ResNet-50	84.30%	85.40%	84.50%	84.95%
EfficientNet	85.20%	85.60%	85.30%	85.45%

4. Analysis and Discussion

As indicated in Table 1, the study compares the performance of ResNet-50 and EfficientNet individually. Both models demonstrate competitive accuracy, precision, recall, and F1-score, highlighting their effectiveness in few-shot clas- sification tasks.

ResNet-50: Achieves an accuracy of 84.30%, with balanced precision, recall, and F1-score around 85%. This indicates ResNet-50's capability in handling few-shot learning scenar- ios. EfficientNet: Shows a slightly higher accuracy of 85.20% compared to ResNet-50. The precision, recall, and F1-score also exhibit balanced performance, reinforcing the suitability of EfficientNet for few-shot image classification.

5. Conclusion

Therefore, both ResNet-50 and EfficientNet are strong can- didates for solving many image classification problems. Their unique architectural features and design make them useful tools for dealing with annotation events. However, a full understanding of their advantages and limitations is important for making informed decisions about their applications in specific areas.

Our results show that EfficientNet outperforms ResNet-50 in several benchmarks, especially when applied to the VGG- Flowers dataset. This finding is consistent with EfficientNet's design concept, which emphasizes achieving high performance standards with limited resources.

EfficientNet's ability to provide better results on the VGG- flowers dataset proves that it can be a good solution for working with limited data. The integration method used by EfficientNet provides high performance in low-budget ap- plications by simultaneously adjusting the depth, width and resolution of the network.

While EfficientNet shows its power in some situations, it is important to remember: the choice between ResNet- 50 and EfficientNet depends on the specific requirements of the project at hand. With its deep and continuous structure, ResNet-50 can perform well in situations where capture is important. On the other hand, the effectiveness of EfficientNet is more evident in fewer areas.

Ultimately choosing the best tool depends on a detailed understanding of the data, computing resources and perfor- mance characteristics. Our study advances this understanding by demonstrating the effectiveness of EfficientNet in analyzing multiple vaccines, giving clinicians and researchers a better understanding of photographic solutions in situations where information is limited.

References

[1] M. Tan, Q. Le, Efficientnet: Rethinking model scaling for convolutional neural networks, In International conference on machine learning. PMLR, (2019) 6105-6114.

[2] K.M. Nguyen, Do, M. V., & Carneiro, G. (2019). Bayesian deep learning for image restoration: Fix the kernel, estimate the noise. In International Conference on Machine Learning.

[3] G. Ravi, A. Beatson, (2019) Learning with fewer labels: Recent progress on few-shot learning, In Advances in Neural Information Processing Systems.

[4] G.J. Gordon, Bayesian deep learning, Nature: Machine Intelligence, 1(1), (2019) 8.

[5] J. Hu, (2020) Meta-learning for semi-supervised few-shot learning. In International Conference on Learning Representations.

[6] M. Sun, (2021). Prototypical meta-learning with adaptive class balances. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.

[7] W. Wang, F. Liu, J. Liu, L. Xiao, Multi-view relation learning for cross-domain few-shot hyperspectral image classification, IEEE Transactions on Geo- science and Remote Sensing, 61(2), (2023) 1428-1444.

[8] F. He, G. Li, L. Si, L. Yan, F. Li, F. Sun, (2023) Proto-typeFormer: Transformer-based pro- totype learning for few-shot image classification. arXiv preprint arXiv:2302.09783. https://doi.org/10.48550/arXiv.2310.03517

[9] R. Wei, (2023) Dual-channel prototype network (DCPN) using transformers and CNNs for efficient few-shot pathology classification. IEEE Journal of Biomedical and Health Informatics.

[10] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei- Fei, (2009) Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition, 248–255. IEEE.

[11] A. Singh, H. Jamali-Rad, (2022) Transductive decoupled variational inference for few-shot classification. arXiv preprint arXiv:2208.10559.

[12] R. Zhang, Z. Wang, Y. Li, J. Wang, Z. Wang, (2023) FewSAR: A Few-shot SAR Image Classification Benchmark. arXiv preprint arXiv:2306.09592.

[13] K. He, X. Zhang, S. Ren, J. Sun, (2016) Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778.

[14] A. Antoniou, A. Storkey, H. Edwards, (2018) How to train your MAML, arXiv preprint arXiv:1810.09502.

Funding

No funding was received for conducting this study.

Conflict of interest

The Authors have no conflicts of interest to declare that they are relevant to the content of this article.

About The License