

Defog Artificial Intelligence Glasses: Neural Networks for the Imperfect Real World

Nilton Rojas

National University of Engineering
nrojasv@uni.pe

Abstract

This research investigates the generalization capabilities of neural networks in deep learning when applied to real-world scenarios where data often contains imperfections, focusing on their adaptability to both noisy and non-noisy scenarios for image retrieval tasks. Our study explores approaches to preserve all available data, regardless of quality, for diverse tasks. The evaluation of results varies per task, due to the ultimate goal of developing a technique to extract relevant information while disregarding noise in the final network design for each specific task. The aim is to enhance accessibility and efficiency of AI across diverse tasks, particularly for individuals or countries with limited resources, lacking access to high-quality data. The dedication is directed towards fostering inclusivity and unlocking the potential of AI for widespread societal benefit.

Introduction

This research focuses on neural network generalization in deep learning, specifically addressing their performance in real-world scenarios, with an emphasis on noisy image data. In my view, it's crucial to develop neural networks that can be deployed reliably in real-life scenarios: where many images exist in imperfect conditions (e.g. shooting angles, irregular illumination, etc.). I aim to explore approaches or architectures that allow neural networks to perform well under these real-world conditions where data is noisy. In the real world, most data, whether images or text, comes with some level of noise or imperfection. For example, in the field of medical imaging, obtaining pristine medical scans is often limited by equipment availability and financial constraints (Trepani et al. 2021). Similarly, in artistic applications, capturing flawless photographs under varying lighting conditions can be a costly endeavor (Pawar et al. 2022). The industrial sector faces challenges when dealing with noisy sensor data due to equipment limitations (Radlak et al. 2020). In the case of images, factors like camera quality and lighting often deteriorate the resultant image. Another type of noise is unrelated to the main object in the picture. For example, if we analyze images containing cats, their background may contain furniture or other objects that are not cats. I am extremely interested in developing

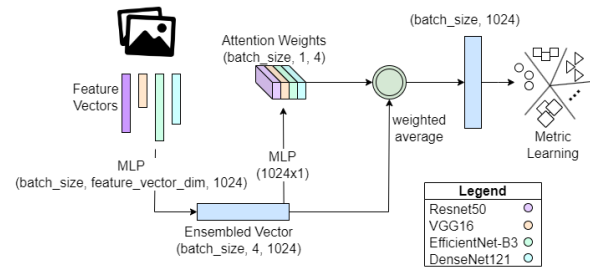


Figure 1. My image-based attention approach with four isolated models

techniques that can handle inputs with a lot of noise. Medical, artistic and industrial societies must reconcile these problems, but take note that it is a general quality issue with image specific data in which I am sure I could help with, given my expertise.

Background

There are some approaches where low-quality images are removed from datasets to avoid noise (Zhang et al., 2020; Liu et al., 2020). However, these approaches often focus on completely removing low-quality data and may not consider the possibility that all the available data is of low quality. Instead of removing all such data, I aim to develop methods which extract knowledge from these low-quality images and adapt it to any task. Currently, I have experimented with configuring neural networks, with the support of various models. The network learns to rely on different models and, even with poor quality inputs, it can extract knowledge and combine different models to perform well. I presented preliminary results of this work at the LatinX in Computer Vision workshop during ICCV 2023 using an attention mechanism for noisy images in an image retrieval task. It generates an intuitive representation of four feature vectors via attention weights and metric learning (figure 1). The data set was collected from the internet in the wild (i.e. with

noise) and has been shared on my personal GitHub account https://github.com/Nicerova7/Noisy_to_Noisy-Free_Retrieval.

Approach

My previous work using an attention mechanism obtains good results when compared to isolated models in image-retrieval task of clothes. However, when analyzing the importance weights of the different models, the results indicate that the contribution of information from all models may be repetitive. We argue that EfficientNet-B3 and DenseNet121 may have redundant features related to ResNet50 (table 1). Most weights are not distributed better across different models. So, these results may only be useful for image classification tasks. In future work, I will use models trained in different tasks such as classification, object detection, segmentation, text-image alignment, multimodal learning, etc., which could help improve performance.

	Resnet50	VGG16	EfficientNet-B3	DenseNet121
Hats	0.118	0.430	0.252	0.200
Tops	0.106	0.489	0.226	0.180
Bottoms	0.101	0.512	0.216	0.171
Shoes	0.105	0.496	0.222	0.178

Table 1. Average weight contribution of each model from our image-based attention mechanism

Evaluation

Generally speaking, the evaluation of my results varies with each task, since the final idea is to have some technique which can extract all the relevant information and ignore the noise. To elaborate, for a classification task it will be an accuracy metric, for an object detection task it will be Average Precision evaluation and so on. The idea is to have well-known models, train them with noisy data and compare the evaluation with or without the technique that I want to perform. Success can be measured by observing an improvement in the evaluation metrics when applying the prefilter or technique.

For now, I am working on an image retrieval task which I am using as a starting point to investigate the best configuration to tackle the problem of noisy inputs. In this task, the evaluation I use is the mean average precision of the first 10 elements (mAP@10). mAP@10 provides a concise measure of how well the system ranks and retrieves relevant items within the top 10 positions. This metric ranges from 0 to 1. Higher values are better.

Discussion

I hope to develop techniques that enhance performance of the models by taking advantage of all types of data. I strongly believe that less developed countries which cannot afford high-quality data, can use my work as a solution so that they can collect data of any quality and use the already known models to be able to obtain useful results in any application regardless of their data quality constraints. Likewise, students who do not have datasets can create their own dataset and also have good results in any task since this is a problem that I and my colleagues encountered where there is a lack of high quality data.

Conclusion

My research focuses on generalizing deep neural networks, specifically addressing the challenge of using all data, whether noisy or low quality. This approach not only seeks to enhance neural networks but also to analyze them for their possible understanding of advantages and limitations in certain domains. Also, we aim to make AI more accessible and efficient for use in different tasks and by people in countries with limited resources who do not necessarily have access to high-quality data. My dedication aims to make AI more inclusive, unlocking its potential for various applications and benefiting society in any given application.

References

- Tripathi, Sumit et al. 2021. An Augmented Deep Learning Network with Noise Suppression Feature for Efficient Segmentation of Magnetic Resonance Images. *IETE Technical Review* 39: 960 - 973. doi.org/10.1080/02564602.2021.1937349.
- Pawar P, Ainapure B, Rashid M, Ahmad N, Alotaibi A, Alshamrani SS. 2022. Deep Learning Approach for the Detection of Noise Type in Ancient Images. *Sustainability*: 14-18. doi.org/10.3390/su141811786.
- Radlak, K., Malinski, L., & Smolka, B. 2020. Deep Learning based Switching Filter for Impulsive Noise Removal in Color Images. *Sensors*, 20(10), 2782. doi.org/10.3390/s20102782.
- Liu, D., Meng, Y., & Wang, L. 2020. Data Cleaning of Irrelevant Images Based on Transfer Learning. In Proceedings of the 2020 International Conference on Intelligent Computing, Automation and Systems (ICICAS), 11-13 December 2020, Chongqing, China. doi.org/10.1109/ICICAS51530.2020.00005.
- Zhang, Y., Jin, Z., Liu, F., Zhu, W., Mu, W., & Wang, W. 2020. ImageDC: Image Data Cleaning Framework Based on Deep Learning. In Proceedings of the 2020 IEEE International Conference on Artificial Intelligence and Information Systems (ICAIS), 20-22 March 2020, Dalian, China. doi.org/10.1109/ICAIS49377.2020.9194788.