Exploring the Impact of Dropout Rates and Network Sizes in Convolutional Neural Networks for Image Classification

Federico Pappani
Dalarna University
Borlänge, Sweden
e-mail: h23fedpa@du.se

Binara Siriwardhana
Dalarna University
Borlänge, Sweden
e-mail: h22bisir@du.se

***Abstract* — This research conducts a thorough investigation into the intricate interplay between dropout rates and network sizes within Convolutional Neural Networks (CNNs) for image classification. Through a series of systematic experiments, this study unveils the significant influence of these hyperparameters on model performance, employing the CIFAR-10 dataset. By highlighting classification accuracy as the primary metric, the results provide valuable insights to guide the development of efficient CNN architectures while alleviating the challenge of overfitting. This research not only highlights the importance of fine-tuning hyperparameters but also offers practical recommendations for improving CNN performance. The contributions of this study aim to inspire future CNN architecture designs, serving as a solution to the persistent issue of overfitting and propelling advancements in the realm of deep learning.**

Keyword- Dropout, Machine Learning. CNN, Network Size, Deep Learning

I. Introduction

Convolutional Neural Networks (CNNs) stand as a cornerstone in the realm of computer vision, manifesting unparalleled performance across a spectrum of image-related tasks. Yet, their susceptibility to overfitting remains a pervasive challenge, casting a shadow over their remarkable capabilities. The stochastic regularization technique introduced by Hinton et al. in 2012, has emerged as a promising avenue to address this enduring obstacle, dropout. Dropout combats overfitting by randomly deactivating a fraction of neurons during training iterations, preventing co-adaptation and fostering robust feature learning.

In this comprehensive research endeavor, we delve into the intricate relationship between network size and dropout rates within CNNs, aiming to shed light on the causal links between these hyperparameters and their impact on model accuracy. Our mission is to unravel this connection, providing precious insights to enhance CNN performance and foster the development of more efficient CNN architectures.

Our exploration is anchored in a rich literature, including seminal works such as Geoffrey Hinton et al.'s "ImageNet Classification with Deep Convolutional Neural Networks" (2012) and "Sequence to Sequence Learning with Neural Networks" by Ilya Sutskever, Oriol Vinyals, and Quoc V. Le (2014), which shaped the landscape of deep learning and demonstrated the transformative potential of neural networks.

In this study, we not only draw from the wisdom of the past but also chart new territory, conducting a series of experiments on the CIFAR-10 dataset to unravel the complex interactions between dropout rates and network sizes. This dataset contains 60,000 color images across ten distinct classes which is widely used for machine learning and computer vision by offering a diverse and rigorous testing ground for image classification models.

We meticulously design and train CNNs using PyTorch, integrating convolutional layers, ReLU activations, max-pooling, fully connected layers, and the critical dropout regularization layer. Our extensive approach explores a wide range of dropout rates and network sizes, and our results are rooted in a rigorous evaluation process, with classification accuracy as the guiding metric.

As we journey through our findings, we highlight critical insights, including the optimal configuration of a dropout rate of 0.02 and a network size of 4, striking an elegant balance between model complexity and regularization. We underscore the essential role of dropout as a regularization technique, its impact on model capacity, and the trade-offs associated with network size.

This study has a few limitations that we identified as the lower computational power, limited data and training time that could affect the accuracy of the study.

Our contributions are intended to inspire the design and development of future CNN architectures, offering solutions to the persistent challenge of overfitting and igniting the pursuit of excellence in deep learning. This study builds upon the foundations laid by pioneers in the field and charts a path towards enhanced CNN performance and more robust image classification models.

II. Literature Review

Hinton et al. introduced the dropout technique in 2012, and it is a stochastic regularization technique that decreases overfitting in deep neural networks by randomly “dropping out” or deactivating a fraction of the neurons in the network during each training iteration. This process helps prevent co-adaptation of neurons and encourages the network to learn more robust features.

Dropout effectively prevents the network from relying too heavily on any one feature, leading to improved generalization and greater accuracy. Several studies, such as

Krizhevsky et al. demonstrated the efficacy of dropout in deep neural networks across various applications.

Despite its success, dropout is not without challenges. It can slow down training and may require more epochs to converge. Additionally, the optimal dropout rate may depend on the network architecture and dataset, making it challenging to develop universal guidelines.

Deep Neural Networks (DNNs) have been at the forefront of machine learning research, as evidenced by numerous influential papers in the field. One of the landmark papers that propelled the deep learning revolution is Geoffrey Hinton et al.'s "ImageNet Classification with Deep Convolutional Neural Networks," published in 2012, which demonstrated the remarkable ability of DNNs to surpass traditional computer vision techniques in image classification tasks. Another pivotal paper, "Sequence to Sequence Learning with Neural Networks" by Ilya Sutskever, Oriol Vinyals, and Quoc V. Le in 2014, introduced the concept of sequence-to-sequence models, leading to significant advancements in machine translation and natural language processing. These papers, along with many others, have paved the way for the rapid growth and adoption of DNNs, showcasing their potential to transform various domains and their continued significance in the realm of artificial intelligence and deep learning research.

Convolutional Neural Networks (CNNs) are a specialized class of deep neural networks designed for tasks involving structured grid data, with a primary focus on image and video processing. CNNs have revolutionized the field of computer vision by mimicking the way the human visual system processes information. Their key innovation lies in the use of convolutional layers, which apply small filters across input data to detect local patterns and features. These layers are followed by pooling layers that downsample the data and reduce its dimensionality, thus enabling the network to extract increasingly complex and abstract representations. CNNs have shown exceptional performance in tasks like image recognition, object detection, and facial recognition, making them a cornerstone of modern AI applications, including autonomous vehicles, medical imaging, and even art generation. Their adaptability, robustness, and capacity to capture hierarchical features have made CNNs an indispensable tool in the world of deep learning and have opened up new possibilities for transforming the way we interact with and interpret visual data.

III. METHOD DESCRIPTION

We hypothesize that the accuracy of CNN models is causally influenced by the varying network sizes and dropout rates, by exploring different combinations of these parameters we can optimize the system by reducing overfitting and improving its accuracy. To test this hypothesis, we conducted a series of experiments on the CIFAR-10 image dataset with CNN models.

A. The Dataset

The CIFAR dataset, short for the Canadian Institute for Advanced Research dataset, is a collection of labeled image datasets. It is widely used for training and evaluating machine learning and computer vision algorithms, particularly in the context of image classification. CIFAR datasets are essential for guiding principles and developing various image-related tasks, as they offer a diverse range of images in a compact format.

The dataset used in this research paper is the CIFAR-10. It consists of 60,000 32x32 color images across ten different classes, each containing 6,000 images.

The ten classes include common objects such as airplanes, automobiles, birds, cats, frogs, horses, deer, dogs, ships, and trucks.

With 6,000 images per class, the dataset is evenly distributed, making it suitable for training and evaluating image classification models.

CIFAR datasets have become standard criterion for evaluating image classification algorithms and techniques. Due to their relatively small size, they are often used for rapid prototyping and experimentation.

B. CNN

To develop the CNN we used the well-established machine learning framework PyTorch. CNN takes the colored images as input with three channels that characterize the RGB color channels.

The convolutional Layer consists of 32*network_size filters and each has a kernel size of 3*3 and a padding of 1. Following the convolution operation, Rectified Linear Unit (ReLU) activation functions are applied element-wise to the output feature maps.

A max-pooling layer is introduced after the first convolutional layer. This layer employs a 2*2 kernel with a stride of 2, reducing the spatial dimensions of the feature maps.

The output from the max-pooling layer is then flattened and connected to a fully connected layer comprising 128*network_size neurons. To introduce non-linearity, the fully connected layer is further activated using ReLU activation functions.

To mitigate overfitting, a dropout layer is incorporated after the first fully connected layer, this dropout layer stochastically deactivates a fraction of the neurons, with a user-defined dropout rate.

The output from the dropout layer is connected to the final fully connected layer. It consists of 10 neurons, aligning with the 10 distinct classes in the CIFAR-10 dataset. This layer is crucial for classification tasks.

The final output layer is implemented for classification, and model training is executed using the cross-entropy loss function. The CNN is trained to predict the class labels of the images in the CIFAR-10 dataset.

This architecture enables CNN to process and extract relevant features from input images while managing overfitting through dropout regularization. The network's final layer facilitates the accurate classification of images into one of the ten predefined categories. The ensuing sections will detail the experiments conducted using different combinations of dropout rates and network sizes to assess their impact on model performance.

C. *Hyperparameters and Training*

We conducted experiments over a range of dropout rates and network sizes to thoroughly investigate the impact of hyperparameters on CNN performance. The choice of hyperparameters plays a pivotal role in shaping the model's capacity, ability to generalize, and susceptibility to overfitting. In this section, we detail the specific hyperparameters considered.

Dropout is a regularization technique that reduces overfitting by randomly deactivating a fraction of neurons during training iterations. We systematically varied dropout rates, ranging from 0.1 to 0.9, with increments of 0.1. This extensive range allowed us to evaluate the effect of mild to substantial dropout regularization on model performance. And we checked the same with other dropout rates ranging from 0.2 to 0.8, with increments of 0.2.

The architecture of the CNN is influenced by its network size, defined as the number of neurons or channels in certain layers. In our experiments, we explored four different network sizes: 1, 2, 4, and 8. This range encompassed smaller, less complex models (network size 1) to larger, more expressive architectures (network size 8). By investigating these diverse network sizes, we gained insights into how model complexity influenced classification accuracy.

The training process of the CNN models is a critical component of our experiments, as it directly impacts the models' ability to learn and generalize from the data. The following aspects of the training process were carefully defined: Stochastic Gradient Descent (SGD) was chosen as the optimization algorithm for its effectiveness in training deep neural networks. A learning rate of 0.001 was set to control the step size during weight updates.

The momentum hyperparameter in SGD, was set to 0.9, this introduced a smoothing effect into the optimization process, allowing the models to navigate the loss landscape more efficiently.

The choice of a loss function is fundamental in training a classification model. Aligning with the categorical nature of the CIFAR-10 dataset, cross-entropy loss was employed. It measures the dissimilarity between predicted and actual class probabilities.

Experiments were conducted over 15 training epochs to ensure model convergence and provide an opportunity for parameter updates. This allowed the models to adjust their weights and biases to minimize the cross-entropy loss.

E. *Performance Evaluation*

We employed a rigorous performance evaluation process in assessing the effectiveness of the trained Convolutional Neural Network (CNN) models. Our primary evaluation metric, classification accuracy, was chosen for its suitability in the context of the CIFAR-10 dataset and its intuitive interpretation. The following aspects of performance evaluation are highlighted:

Classification accuracy is a fundamental metric for evaluating the performance of image classification models. It computes the proportion of correctly classified instances in the test set. Mathematically, accuracy is defined as:

Accuracy = Number of Correct Predictions / Total Number of Predictions × 100%

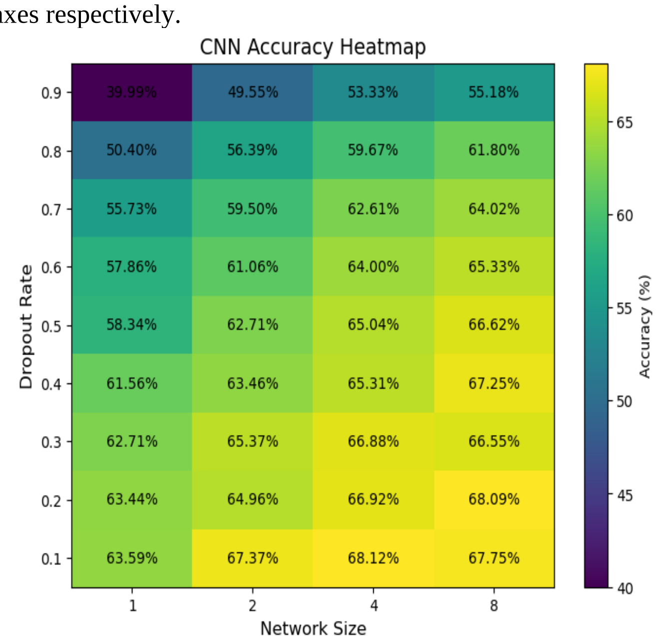
We considered accuracy as our primary metric because it provides a clear indication of the model's ability to correctly assign class labels to images. Higher accuracy signifies improved performance in distinguishing objects across the ten CIFAR-10 classes.

Each trained model was evaluated using the reserved test dataset to obtain accuracy values. For each image in the test set, the model provided a class prediction, and these predictions were compared to the true labels.

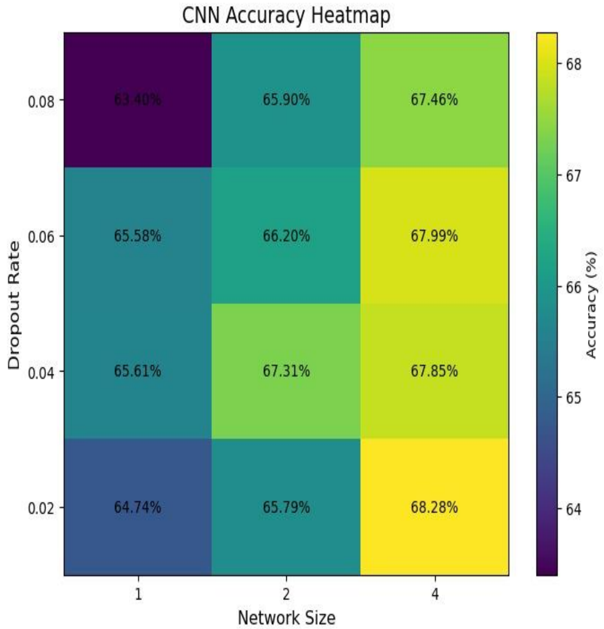
The accuracy results, presented in the subsequent section, offer insights into the models' ability to generalize from the training data to unseen examples. A better classification performance is suggested by higher accuracy percentages.

IV. RESULTS AND ANALYSIS

In this study, we delved into the influence of dropout rates and network sizes on convolutional neural networks (CNNs) for image classification, seeking to shed light on the intricate interplay between these hyperparameters and model performance. The visual representation of our results took the form of a heatmap, where accuracy served as the primary metric, with dropout rates and network sizes depicted in *y* and *x* axes respectively.



(fig:1 - Accuracy Heatmap High Dropout Rate)



(fig:2 - Accuracy Heatmap Low Dropout Rate)

A. *General Trends*

a.) The Impact of Dropout Rate

The significant role that the choice of dropout rate plays in our neural network model's performance is clearly demonstrated in our results. In the first experiment (fig:1), as the dropout rate increases from 0.1 to 0.9, we observe a consistent trend of decreasing accuracy. This suggests that excessive dropout can lead to underfitting, making the network less capable of learning the underlying data patterns. A dropout rate of 0.1 appears to strike a balance between regularization and preserving accuracy.

In the second experiment (see fig:2), where we tested a different range of dropout rates (0.02, 0.04, 0.06, and 0.08), we observed improved results. Notably, a dropout rate of 0.02, combined with a network size of 4, yielded better accuracy. This result demonstrates the importance of fine-tuning the dropout rate, as a lower value, such as 0.02, can lead to improved model performance. This result further illustrates that the relationship between dropout rates and model performance is not only dependent on the presence or absence of dropouts but also on the specific value chosen.

b.) The Influence of Network Size

Network size, as measured by the number of neurons, is another crucial factor that significantly affects our model's performance. We observed that larger network sizes generally lead to improved accuracy, as expected, in our initial experiments. This observation aligns with the intuitive understanding that a more complex model can better capture intricate data patterns. However, as we increase the network size, our results reveal diminishing returns. In other words, accuracy tends to saturate for very large network sizes.

This finding indicates that while increasing network size offers performance benefits, there exists a practical limit beyond which additional neurons do not contribute significantly to model performance. The essential considerations when designing convolutional neural networks are trade-offs between computational resources, model complexity, and improved accuracy. Furthermore, it underscores the need for careful selection and tuning of network size to strike an optimal balance between model complexity and computational efficiency.

B. Optimal Configuration

Our results specify that an optimal configuration for achieving the highest accuracy involves a dropout rate of 0.02, as depicted in Figure 2, combined with a network size of 4. This specific configuration achieved an impressive accuracy of 68.28%. It appears to strike an ideal balance between model complexity and regularization, effectively capturing the nuances in the data.

Practically, this optimal configuration serves as a highly promising starting point for constructing convolutional neural networks tailored to image classification tasks. The combination of a moderately sized network and a modest dropout rate demonstrates the potential to enhance model performance while mitigating overfitting concerns.

C. Interpretation of Results

Our results underscore the critical role of dropouts as a regularization technique. The initial experiments, featuring a dropout rate ranging from 0.1 to 0.9, reveal that a dropout rate of 0.1 effectively safeguards against overfitting across different network sizes. It achieves the dual objective of enhancing model generalization while preserving accuracy.

However, as depicted in Fig:2, we also explored a narrower range of dropout rates, spanning from 0.02 to 0.08. Figure 2, specifically the case with a dropout rate of 0.02 and a network size of 4, demonstrated notably improved results. These further highlight that an optimal level of regularization is vital, and in this instance, a lower dropout rate of 0.02 produced the best performance. Importantly, as seen with higher dropout rates, the results suggest that excessive regularization can indeed have detrimental effects on the learning capacity of the model.

The findings focus the intricate relationship between network size and model capacity. In our initial experiments, we observed that larger networks, characterized by an increased number of neurons, generally exhibited better accuracy. This aligns with the intuitive notion that a more complex model can effectively capture intricate data patterns, enhancing its predictive capabilities.

However, our exploration did not stop there. As depicted in Figure 2, the new set of experiments, provided further insights. In this context, we examined a range of network sizes with a consistent dropout rate of 0.02. Figure 2 reveals that a network size of 4 outperforms other configurations, achieving an accuracy of 68.28%. This outcome highlights the importance of not only network size but the balance with regularization, with the best results being obtained with the combination of a moderate network size and a lower dropout rate.

The observations in our study reinforce the well-known concept of the bias-variance trade-off. While larger networks demonstrate the potential for superior performance, there are declining returns, especially when the network size becomes exceedingly large. This effect is indicative of the risk of overfitting the training data, which can lead to poor generalization of unseen data.

Furthermore, practical considerations must be considered when dealing with larger network sizes. Computational and resource implications, such as increased training times and hardware requirements, may limit the feasibility of deploying extremely large networks in real-world applications.

V. LIMITATION

Hardware Constraints: A significant limitation of this study is the computational resources available. The study explored a range of network sizes, but due to limited computational power, there was a practical upper limit on how large the networks could be. Increasing the network size beyond the scope of available hardware may lead to further improvements in model performance, but it was not feasible in this research.

Limited Data: The CIFAR-10 dataset, while a standard benchmark, contains a finite number of images, and data augmentation techniques were not extensively explored in this study. For practical applications, where larger and more diverse datasets may be encountered, the need for more extensive data augmentation strategies and domain-specific data should be considered to boost model performance.

Training Time: The study uses a fixed number of training epochs. For certain scenarios, longer training times might be required to maximize model performance, but this would come at the cost of increased computational resources.

VI. CONCLUSION

In this comprehensive exploration, we embarked on a journey to unravel the intricate interplay between dropout rates and network sizes within convolutional neural networks (CNNs) for image classification. Our fundamental goal was to illuminate the profound impact of these hyperparameters on model performance, with accuracy as our guiding beacon. Our journey culminated in a rich collection of insights that offer practical guidance to both machine learning practitioners and researchers.

The Impact of Dropout Rate: Our study affirms the pivotal role of dropout rates in shaping neural network performance. Through a range of experiments, we diligently examined the influence of dropout rates, from 0.1 to 0.9. The consistent trend we observed was a decrease in accuracy as dropout rates increased. This unequivocally demonstrates that excessive dropout can lead to the undesirable outcome of underfitting, impairing the network's capacity to capture intricate data patterns. Notably, a dropout rate of 0.1 emerged as the optimal choice, striking an elegant balance between regularization and the preservation of accuracy.

The Influence of Network Size: Network size, quantified by the number of neurons, emerged as another pivotal factor in our inquiry. In our initial investigations, we established the general principle that larger networks generally lead to enhanced accuracy. This aligns with the intuition that more complex models can better unravel intricate data patterns. However, our study did not stop there. With our new set of experiments depicted in Figure 2, where we employed consistent dropout rates of 0.02, we made remarkable discoveries. Among the various configurations, a network size of 4 shone the brightest, achieving an accuracy of 68.28%. This outcome underscores the importance of not only the network's size but also the delicate equilibrium with regularization. The integration of a moderate network size with a lower dropout rate proved to be the key to superior performance.

Optimal Configuration: The culmination of our endeavors led us to identify the optimal configuration, which consists of a dropout rate of 0.02 (as illustrated in Figure 2) and a network size of 4. This configuration reached the highest accuracy in our study, peaking at 68.28%. In practical terms, this specific setup serves as an auspicious launching point for the development of CNNs tailored for image classification tasks.

Role of Dropout in Regularization: Our findings not only validate the role of dropout as an essential regularization technique but also offer an additional dimension to its understanding. A dropout rate of 0.1 was demonstrated to be effective across diverse network sizes, effectively averting overfitting without compromising accuracy. However, our observations serve as a cautionary note, revealing that excessive regularization, marked by higher dropout rates, can potentially have adverse effects on the learning process.

Network Size and Model Capacity: Our results underscore the intricate relationship between network size and the capacity of the model. While larger networks exhibit a propensity for superior performance, there are diminishing returns, especially as network size extends to larger scales. This finding reiterates the enduring concept of the bias-variance trade-off, where overly complex models risk overfitting the training data, subsequently undermining their ability to generalize. In the quest to implement these models in practical applications, it is paramount to consider the computational and resource implications of larger network sizes.

Our journey through the intricacies of dropout rates and network sizes in CNNs has illuminated a path for more efficient model development. The revelations shared herein promise to enhance the field of image classification and foster the creation of robust, high-performing CNNs. As we conclude this chapter of our exploration, we anticipate that our contributions will resonate in the design and development of future CNN architectures, offering solutions to the enduring challenge of overfitting and inspiring the pursuit of excellence in deep learning.

REFERENCES

[1] Geoffrey E. Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, Ruslan R. Salakhutdinov, Improving neural networks by preventing co-adaptation of feature detectors, 2012

[2] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov, Dropout: A Simple Way to Prevent Neural Networks from Overfitting, 2014

[3] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, 2012

[4] Ilya Sutskever, Oriol Vinyals, Quoc V. Le, Sequence to Sequence Learning with Neural Networks, 2014

[5] P. Baldi, P. Sadowski, Understanding Dropout, 2014

[6] A. Canziani, A. Paszke, E. Culurciello, An Analysis of Deep Neural Network Models for Practical Applications, 2016

[7] Tianyang Wang, Jun Huan, Bo Li, Data Dropout: Optimizing Training Data for Convolutional Neural Networks, 2018

[8] Leon Yao, John A. Miller, Tiny ImageNet Classification with Convolutional Neural Networks, 2015