**CNN Model Performance Analysis Report**

**Assignment Overview**

Deep learning models, particularly Convolutional Neural Networks (CNNs), have demonstrated significant advancements in image classification tasks. These models leverage multiple layers to detect patterns and extract meaningful features from images. However, their performance is heavily influenced by factors such as dataset size and the implementation of regularization techniques. Overfitting, underfitting, and generalization are common challenges in training CNN models, making it essential to explore various strategies to optimize performance.

This assignment aims to investigate the impact of dataset size and regularization methods on CNN model performance by training multiple configurations and evaluating their effectiveness. The study begins with training CNNs from scratch using datasets of varying sizes and implementing regularization techniques like dropout and L2 regularization to understand their impact on accuracy and overfitting. Additionally, a pre-trained VGG16 model is fine-tuned to compare its performance with custom-trained CNNs.

By systematically increasing the training dataset size and adjusting regularization strategies, this study seeks to determine the most effective approach for achieving high classification accuracy while mitigating overfitting. The findings provide valuable insights into best practices for training CNN models and highlight the trade-offs between dataset size, regularization methods, and model complexity. Furthermore, the comparison between training CNNs from scratch versus leveraging transfer learning showcases the advantages of using pre-trained networks to enhance classification performance.

**1. Overfitting Reduction & Model Evaluation**

**1.1 Baseline Model (No Regularization)**

The baseline model serves as the foundation for comparison. This CNN was trained without any regularization techniques to observe how well it performs on the dataset without constraints.

* **Training Accuracy:** 0.95
* **Validation Accuracy:** 0.75
* **Testing Accuracy:** 0.702
* **Observation:** The model exhibits clear signs of overfitting. While it achieves a high training accuracy of 95%, its test accuracy drops to 70.2%, indicating that the model memorized the training data rather than learning generalizable patterns.

**1.2 Dropout Regularization (0.5 Rate)**

Dropout is a widely used technique to prevent overfitting by randomly deactivating a fraction of neurons during training.

* **Training Accuracy:** 0.97
* **Validation Accuracy:** 0.75
* **Testing Accuracy:** 0.497
* **Observation:** While dropout successfully reduces overfitting, its aggressive application led to underfitting, causing a significant drop in test accuracy. This suggests that dropout was too high, preventing the model from learning key features.

**1.3 L2 Regularization**

L2 regularization penalizes large weight values to encourage simpler models with better generalization capabilities.

* **Training Accuracy:** 0.49
* **Validation Accuracy:** 0.50
* **Testing Accuracy:** 0.500
* **Observation:** The model suffered from severe underfitting due to excessive weight penalties, restricting its ability to learn meaningful features.

**1.4 Combined Dropout & L2 Regularization**

This approach combines the benefits of both dropout and L2 regularization to balance overfitting and underfitting.

* **Training Accuracy:** 0.69
* **Validation Accuracy:** 0.64
* **Testing Accuracy:** 0.685
* **Observation:** While this approach controlled overfitting better than using dropout or L2 alone, it did not surpass the baseline model in overall accuracy.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Training Samples** | **Method** | **Training Acc** | **Validation Acc** | **Testing Acc** |
| Baseline | **1000** | None | **0.95** | **0.75** | **0.702** |
| Dropout | **1000** | Dropout (0.5) | **0.97** | **0.75** | **0.497** |
| L2 | **1000** | L2 | **0.49** | **0.50** | **0.500** |
| Dropout + L2 | **1000** | Dropout (0.25) + L2 | **0.69** | **0.69** | **0.684** |

A graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of

AI-generated content may be incorrect.

A graph of a graph with blue dots

AI-generated content may be incorrect.

A screenshot of a graph

AI-generated content may be incorrect.

A screenshot of a graph

AI-generated content may be incorrect.

**Observations & Results**

The baseline model performed well during training, reaching 95% accuracy. However, overfitting was evident as testing accuracy dropped significantly to 70.2%. This indicates that the model memorized the training data instead of learning generalizable patterns. The validation accuracy stagnated at 75%, further confirming limited generalization ability. While dropout helped reduce overfitting, an aggressive dropout rate of 0.5 led to underfitting. The training accuracy increased to 97%, but test accuracy declined to 49.7%. This suggests that the model failed to learn essential patterns due to excessive dropout, impairing its ability to generalize to unseen data.

Applying L2 regularization significantly restricted the model’s learning capability, leading to severe underfitting. The training accuracy dropped to 49%, and validation and test accuracies remained around 50%, indicating that the model performed no better than random chance. This demonstrates that excessive regularization can limit the model’s capacity to extract meaningful features. The combined approach improved generalization more effectively than dropout or L2 alone. The training accuracy of 69% suggests that the model was regularized effectively without extreme underfitting. The test accuracy of 68.5% was higher than the L2-only and dropout-only models, demonstrating better generalization. However, performance remained lower than the baseline model, indicating the need for optimized regularization strategies.

**2. Impact of Training Set Size on Accuracy**

The number of training samples plays a crucial role in improving CNN performance. Increasing dataset size often enhances model generalization and reduces overfitting.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Training Samples | Method | Training Accuracy | Training Accuracy | Testing Accuracy |
| Baseline | **1000** | None | **0.95** | **0.75** |  |
| Model 1500 | **1500** | None | **0.98** | **0.76** | **0.716** |
| Dropout 2000 | **2000** | Dropout (0.5) | **0.95** | **0.76** | **0.728** |

**Key Observations:**

* The 1000-sample model showed clear overfitting.
* Increasing to 1500 samples improved accuracy slightly.
* Training with 2000 samples and applying dropout led to the highest test accuracy (72.8%), confirming the benefits of larger datasets.

**Observations & Results:**

* The model trained with 1000 samples achieved high training accuracy, but overfitting was apparent as test accuracy remained lower.
* Increasing the dataset size to 1500 samples improved training and validation accuracy slightly, with test accuracy increasing from 70.2% to 71.6%, suggesting improved generalization.
* The model trained on 2000 samples with dropout achieved the highest test accuracy of 72.8%. This indicates that a larger dataset, combined with regularization, allowed the model to generalize more effectively.
* While increasing training data led to performance improvements, the accuracy gains started to plateau beyond 1500 samples, suggesting that additional techniques such as fine-tuning or deeper architectures may be necessary for further improvements.

A graph with blue dots

AI-generated content may be incorrect.

A graph with blue lines and dots

AI-generated content may be incorrect.

**3. Transfer Learning with VGG16**

Transfer learning leverages pre-trained models to improve accuracy and training efficiency.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Training Samples | Methods | Training Acc | Validation Acc | Test Acc |
| Basic CNN | 1000 | None | 0.95 | 0.75 | 0.702 |
| Basic CNN | 500 | None | 0.95 | 0.76 | 0.722 |
| Basic CNN | 1500 | None | 0.98 | 0.76 | 0.716 |
| CNN + Dropout | 1000 | Dropout (0.5) | 0.97 | 0,75 | 0.497 |
| CNN + Dropout | 2000 | Dropout (0.5) | 0.95 | 0,76 | 0.728 |
| CNN + L2 Regularization | 1000 | |  | | --- | |  |  |  | | --- | | L2 | | 0.49 | 0.50 | 0.500 |
| CNN + Dropout + L2 | 1000 | |  | | --- | |  |  |  | | --- | | Dropout (0.25) + L2 | | 0.69 | 0.64 | 0.685 |
| Feature Extraction (VGG16) | N/A | |  | | --- | |  |  |  | | --- | | Dropout | | 0.99 | 0.97 | 0.977 |
| Fine-tuned VGG16 | 2000 | Dropout + Augmentation | 0.99 | 0.97 | 0.976 |

A graph with blue squares

AI-generated content may be incorrect.

**Refined Conclusion**

This study highlights the significant impact of dataset size, regularization techniques, and transfer learning on CNN model performance. Each factor contributes to the model’s ability to generalize effectively and achieve high classification accuracy. Below is a detailed summary of the key findings:

**1. Larger Dataset Size Enhances Model Generalization**

Expanding the training dataset led to noticeable improvements in model performance. With only 1000 training samples, the model exhibited moderate overfitting, as indicated by the substantial gap between training accuracy (95%) and test accuracy (70.2%). However, increasing the dataset size to 1500 and 2000 samples enhanced test accuracy to 71.6% and 72.8%, respectively. This demonstrates that a larger dataset allows the model to recognize patterns more effectively, reducing its reliance on memorization.

**Key insights:**

* More data exposes the model to diverse features, improving its ability to generalize.
* Overfitting decreases as the model learns broader patterns rather than memorizing training examples.
* Performance gains plateaued beyond 1500 samples, suggesting that additional strategies, such as transfer learning, may be needed for further improvements.

**2. Balanced Regularization is Crucial to Prevent Underfitting and Overfitting**

Regularization methods such as dropout and L2 weight penalties play a key role in controlling overfitting, but their effectiveness depends on appropriate tuning. A dropout rate of 0.5 led to significant underfitting, reducing test accuracy to 49.7%, while excessive L2 regularization restricted the model’s learning, resulting in an accuracy of only 50%.

However, combining moderate dropout (0.25) with L2 regularization led to improved generalization, achieving a test accuracy of 68.5%. This indicates that while regularization is essential, applying it too aggressively can hinder the model’s ability to extract meaningful features.

**Key insights:**

* Excessive regularization limits learning and causes underfitting.
* Insufficient regularization leads to overfitting, where the model memorizes training data rather than identifying general patterns.
* A balanced approach combining dropout and L2 regularization yields better generalization compared to using either technique in isolation.

**3. Transfer Learning with VGG16 Delivered the Best Accuracy**

Among all models tested, the highest test accuracy (97.6%) was achieved using transfer learning with the VGG16 architecture. Fine-tuning the pre-trained model, along with data augmentation, significantly outperformed all CNNs trained from scratch.

Since VGG16 has been pre-trained on large datasets like ImageNet, it already captures essential visual features, allowing it to adapt to new classification tasks with minimal training. This approach proves to be more efficient and effective than training a model from the ground up.

**Key insights:**

* Feature extraction from VGG16 enabled the model to achieve high accuracy with minimal training.
* Fine-tuning the top layers of the pre-trained model further enhanced its classification performance.
* Compared to building CNNs from scratch, transfer learning provided a clear advantage, achieving superior accuracy in less time and with fewer data samples.

**Final Takeaway**

For optimal CNN performance, a well-balanced combination of dataset size, regularization, and transfer learning is essential:

* Increasing dataset size improves generalization but has diminishing returns beyond a certain threshold.
* Regularization should be applied carefully to avoid over-constraining the model or allowing it to overfit.
* Transfer learning with pre-trained models, such as VGG16, is the most effective approach, delivering high accuracy with minimal training effort.

By leveraging these strategies, CNN models can be optimized for robust performance, ensuring high accuracy while minimizing overfitting in real-world image classification tasks.

**Final Recommendation**

For optimal image classification, using transfer learning with a pre-trained model, combined with moderate regularization and data augmentation, is recommended for achieving high accuracy while minimizing overfitting.