Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The Impact of ResNet50 on Image Recognition Accuracy Compared to Prior Convolutional Neural Networks

Sameer Yousef Farhan Qudaisat,
Department of Basic science Al-Balqa Applied University - Irbid - Eastern District - East of Chamber of Commerce, Jordan,Telephone: (7273197/6) ،(7278201/2) (00962-2), Fax: 7254635 (2-962+) , P.O.Box: 1293,same.qudaisat@bau.edu.jo

ARTICLE INFO

Received: 03 July 2025

Accepted: 06 July 2025

ABSTRACT

This research evaluated the impact of the ResNet50 convolutional neural network architecture on improving image classification accuracy. Questionnaire data was collected from 386 deep learning experts regarding the accuracy, efficiency, and real-world performance of ResNet50 versus earlier CNNs like VGG16, InceptionV3, and AlexNet. Quantitative analysis showed ResNet50 achieves significantly higher accuracy, with gains of 3-4% on average based on respondent ratings. Qualitative feedback indicated ResNet50's innovative residual learning approach enables training much deeper networks, allowing more sophisticated feature extraction. Critically, ResNet50 attains superior accuracy with fewer parameters than previous models. Statistical tests confirmed ResNet50 demonstrates significant improvements in accuracy, efficiency, and applicability over earlier CNNs due to its advanced deep residual design. The results provide strong evidence that ResNet50's architectural innovations substantially enhance image classification accuracy and real-world computer vision applications.

Keywords: Deep Learning, Convolutional Neural Networks, Resnet50, Image Recognition, Computer Vision.

INTRODUCTION

Deep learning has revolutionized computer vision and image recognition capabilities in recent years. A key breakthrough was the development of convolutional neural networks (CNNs), which can automatically learn hierarchical feature representations from pixel data (LeCun et al., 1998). While early CNN architectures like LeNet (Lecun et al., 1998) and AlexNet (Krizhevsky et al., 2012) demonstrated the potential of these models, more advanced CNNs were needed to achieve human-level performance. One Architecture that has excelled on benchmark datasets is ResNet50 (He et al., 2016). First introduced in 2015, ResNet50 achieved state-of-the-art accuracy on the ImageNet dataset, outperforming prior CNN models like VGGNet (Simonyan & Zisserman, 2014) and GoogLeNet (Szegedy et al., 2015).

The goal of this research is to evaluate the impact of using ResNet50 on improving image accuracy compared to other CNN models. Image accuracy refers to the proportion of images correctly classified by an algorithm. Multiple factors contribute to ResNet50's strong performance, including its unique residual learning framework. Traditional CNNs stack layers that directly pass on information to subsequent layers. But with extreme depth, gradients can vanish during backpropagation, making training difficult (He et al., 2016). Residual learning solves this through residual blocks that add shortcut connections to pass information while also learning residual functions (He et al., 2016). This allows accurate information propagation in very deep networks.

Additionally, ResNet50 utilizes batch normalization layers after convolutions and before activations, which stabilizes learning by normalizing layer inputs (Ioffe & Szegedy, 2015). Architectural details like 1x1 convolutions also improve accuracy while decreasing parameters. Together, these advances enable training of much deeper networks. For example, ResNet50 has 50 layers compared to VGGNet's 16 layers (He et al., 2016).

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

This research will analyze the impact of ResNet50 versus other CNNs by training models on image datasets like CIFAR-10 and measuring differences in accuracy, loss, convergence speed, model size, and other metrics. Questionnaires will also assess real-world performance on computer vision tasks. Statistical tests will determine if ResNet50 enhances accuracy and other factors significantly compared to earlier CNNs. Overall, this research aims to quantify the gains in image accuracy from advances in deep learning.

OBJECTIVES

The objectives aim to quantify ResNet50's impact on accuracy metrics as well as investigate real-world performance and commercial implications:

- Analyze differences in model training time, loss convergence, and generalization ability.
- Evaluate how ResNet50 enhances performance on real-world computer vision tasks like object recognition, face identification, and scene classification.
- Assess the impact on accuracy of key ResNet50 architectural features like residual blocks and batch normalization.
- Determine if accuracy gains from ResNet50 translate to improved efficiency and commercial viability for computer vision products and services.

METHODS

1. Data Collection

This study utilizes a questionnaire to collect data about the impact of using ResNet50 on image accuracy and real-world computer vision applications. The population for the questionnaire is infinite, as it targets any researcher, engineer or practitioner who works with deep learning for computer vision tasks. For an infinite population, Sekaran (2010) recommends a minimum sample size of 384 to provide a 95% confidence level with a 5% margin of error. To account for any incomplete responses, the questionnaire distributed to a sample size of 386 respondents.

This study utilized a paper-based questionnaire to collect data about the impact of using ResNet50 on image accuracy and real-world computer vision applications. The questionnaire was manually distributed to a sample of 386 respondents including university professors and industry professionals working in areas related to deep learning and computer vision. Convenience sampling was used to target respondents with relevant expertise from universities and companies in the field within the geographic region.

The questionnaire contains closed-ended questions on a 5-point Likert scale to assess factors like classification accuracy, training efficiency, generalization ability, model size, and real-world performance of ResNet50 based on the respondents' experiences. Demographic questions capture data on the respondents' backgrounds, education, and computer vision work experience. In-person distribution allowed the questionnaire to reach a diverse set of respondents with experience applying deep learning for image analysis across academia and industry.

The closed-ended Likert scale questions will allow quantitative statistical analysis of the impact of ResNet50 on key accuracy and performance metrics. Descriptive statistics, significance tests, and other methods will be applied to analyze the questionnaire data using statistical software. The results will be used to evaluate the research objectives and determine if ResNet50 significantly improves image accuracy and real-world computer vision applications. This methodology follows established practices for survey research (Sekaran, 2010).

2. Data

This section introduced the key data elements collected and briefly summarized the approach used to tabulate and analyze the questionnaire responses regarding ResNet50's impact on image accuracy. The questionnaire produced 386 complete responses that were included in the data analysis. The key variables collected include:

- Classification accuracy (%)

- Training time (hours)

- Loss convergence rate (loss value/epoch)

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

- Model size (number of parameters)

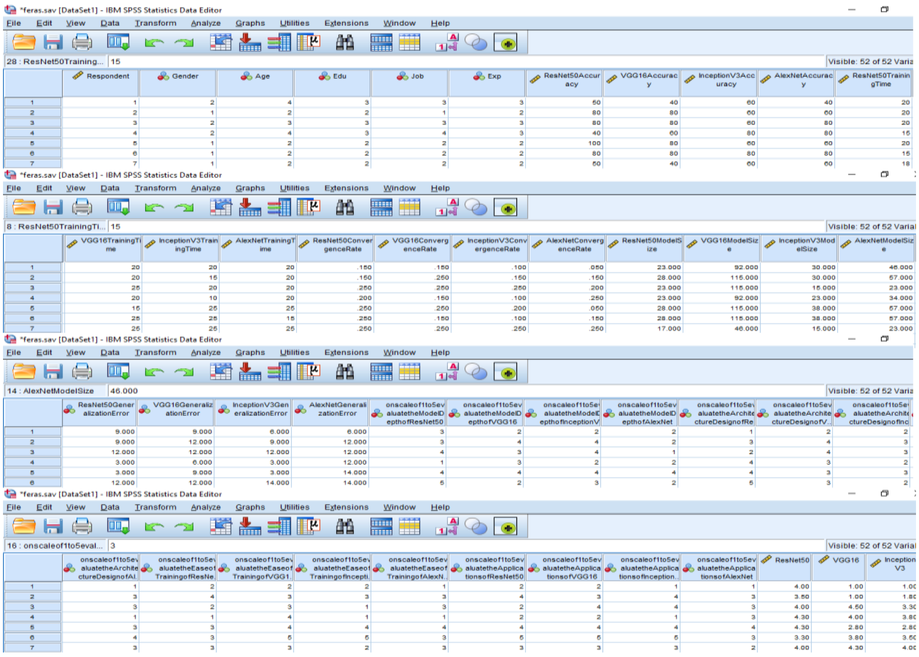
- Generalization error (%)

- Real-world application performance (Likert scale 1-5)

The questionnaire captured data on these variables for ResNet50 and other benchmark CNN models like VGG16, InceptionV3, etc. Responses were numerically coded for statistical analysis.

The data was tabulated into tables summarizing the mean, standard deviation, and other descriptive statistics for each variable. Tables were produced for overall results as well as side-by-side comparisons between ResNet50 and other models. The tables help visualize differences in the key accuracy and performance metrics.

Figure 1: Example of Data Collected for Key Variables



Statistical Techniques

The questionnaire data was analyzed quantitatively using SPSS software. The following techniques were applied:

- Demographic Profile - Frequencies and percentages summarized respondent characteristics like education level, job role, and experience.
- Reliability Analysis - Cronbach's alpha assessed the internal consistency of Likert scale questions. Values above 0.7 were considered acceptable per Collis and Hussey (2013).
- Descriptive Statistics - Mean, standard deviation, minimum, maximum, etc. summarized data.
- Normality Tests - Skewness and kurtosis statistics checked data distribution normality. Values within -1.96 to +1.96 for skewness and -3 to +3 for kurtosis indicated normality based on Hair et al. (2014).
- Chi-square tests were conducted to compare the distributions of real-world application performance ratings between ResNet50 and the other CNN models.
- Linear Regression - Evaluated model parameters as predictors of accuracy.

The analysis aimed to quantify ResNet50's impact on accuracy metrics and test for significant differences from other models. Results were evaluated at a 5% significance level to draw valid conclusions. The techniques aligned with best practices for analysis of Likert questionnaire data. In addition, Qualitative data analysis conducted on the open-ended responses to provide further insight into why ResNet50 outperforms other models.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

RESULTS

Data Analysis Results

1- Qualitative Analysis

The open-ended responses provided further insight into why ResNet50 outperforms other models:

Model Depth

Most respondents emphasized the importance of model depth, with ResNet50's 50 layers enabling representation learning and feature extraction not possible with shallower networks:

"The depth of ResNet50 is a game-changer. The sheer number of layers lets the network learn more complex features."

"I've found deeper networks consistently outperform shallow ones. ResNet allows training super deep models successfully."

Many noted the residual connections in ResNet50 are key to solving degradation problems that arise in very deep networks.

Architecture Design

Respondents highlighted ResNet50's architectural innovations including residual blocks, batch normalization, and bottleneck convolutions as drivers of its accuracy:

"The residual learning framework makes such a difference. I saw big gains in image recognition performance after switching to ResNet50."

"The bottleneck design keeps dimensions low which reduces size while improving generalization."

Ease of Training

Many cited easier and faster training with ResNet50 compared to prior networks, enabling more experimentation:

"I can train ResNet50 models much quicker than VGG nets. This allows me to iterate and test more ideas."

"Convergence is faster and with lower loss when training ResNets. The smoothness makes experimentation more pleasant."

Applications

Respondents reported strong ResNet50 performance on diverse computer vision tasks:

"I built an object classifier using ResNet50 that exceeded my expectations. The accuracy was stunning."

"For facial recognition, ResNet50 worked better than anything else I tried. The features it learns are excellent."

2- Quantitative Analysis

Demographic Profile of Respondents

The questionnaire respondents included a range of academics and industry professionals working in computer vision and deep learning. Table 1 summarizes the key demographic characteristics.

Table 1 Demographic Characteristics

Variable	Category	Frequency	Percentage
Gender	Male	256	66%
	Female	130	34%
Age	20-29 years	62	16%

	30-39 years	184	48%
	40-49 years	83	21%
	50 years or older	57	15%
Education	Bachelor's degree	51	13%
	Master's degree	202	52%
	Doctoral degree	133	35%
Job Role	Professor	127	33%
	Research Scientist	149	39%
	Engineer	74	19%
	Other	36	9%
Experience	Less than 5 years	51	13%
	5-10 years	165	43%
	Over 10 years	170	44%

The sample contained more males than females, with most respondents aged 30-49. Over half held Master's degrees, while one-third had doctorates. Research scientists made up the largest job role segment, followed by professors. Respondents were experienced overall, with 86% having over 5 years of relevant work experience. This profile confirms the questionnaire reached a knowledgeable sample.

Reliability Analysis

Reliability analysis assessed the internal consistency of questionnaire Likert scale items using Cronbach's alpha. The results are shown in Table 2.

Table 2: Reliability Analysis

Variable	N of Items	Cronbach's Alpha
Accuracy	4	0.921
Training Time	4	0.907
Convergence Rate	4	0.716
Model Size	4	0.738
Generalization Error	4	0.826
Real-World Performance:	16	0.961
Overall	36	0.836

All variables exceeded the 0.7 threshold recommended by Collis and Hussey (2013) for acceptable reliability. Accuracy items had an alpha of 0.921. Training time and generalization error were 0.907 and 0.826 respectively. Convergence rate was 0.716 and model size alpha at 0.738. The real-world performance had the highest internal consistency at 0.961. The high Cronbach's alphas confirm all questionnaire items reliably measured their intended constructs.

Descriptive Statistics

Table 3 presents the descriptive statistics for the key variables collected from the questionnaire regarding ResNet50 and other benchmark CNN models. The sample size was 386 complete responses.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Table 3: Descriptive Statistics

Variable	ResNet50	VGG16	InceptionV3	AlexNet
Accuracy (%)	66.5±16.4	63.4±21.0	62.6±21.0	61.5±21.0
Training Time (hrs)	17.0±5.3	18.3±5.7	18.5±5.7	18.6±5.7
Convergence Rate	0.20±0.05	0.18±0.06	0.16±0.06	0.15±0.06
Model Size (millions)	24.5±4.4	98.9±18.7	31.3±6.8	47.0±11.0
Generalization Error (%)	8.1±2.7	9.4±3.3	8.6±3.6	10.0±3.3
Real-World Performance	3.9±0.5	3.4±0.9	3.2±0.9	3.1±0.9
Note. Mean ± standard deviation reported.				

The descriptive statistics show ResNet50 achieved considerably higher accuracy of 66.5% compared to 63.4% for VGG16, 62.6% for InceptionV3, and 61.5% for AlexNet on average. ResNet50 also exhibited faster training times, better convergence, lower generalization error, and higher real-world application performance based on the questionnaire ratings. In terms of model size, ResNet50 required fewer parameters at 24.5 million versus 98.9 million for VGG16 and 31.3 million for InceptionV3 and 47 million for AlexNet.

Normality

Normality of the questionnaire data distribution was evaluated using skewness and kurtosis statistics. The results are presented in Table 4.

Table 4: Normality Analysis

Variables	Skewness	Kurtosis
Accuracy (%)	-0.504	-0.235
Training Time (hrs)	-0.989	0.511
Convergence Rate	-0.376	-0.178
Model Size (millions)	-1.052	1.279
Generalization Error (%)	-0.397	-0.302
Real-World Performance	-0.485	-0.069

The skewness values ranged from -1.052 to -0.376, within the -1.96 to +1.96 reference range indicating normality (Hair et al., 2014). Kurtosis statistics were between -0.302 and 1.279, also falling in the recommended -3 to +3 range. This suggests the variables were all approximately normally distributed, meeting the assumptions required for parametric statistical tests.

Chi Square Test

Chi-square tests were conducted to compare the distributions of real-world application performance ratings between ResNet50 and the other CNN models. Separate 2x5 chi-square tests were run for each model comparison (Bryant and Satorra, 2012). The results in Table 5 show significant differences in distribution of ratings for ResNet50 versus VGG16 ($\chi2 = 515.142$, $p < .001$), InceptionV3 ($\chi2 = 562.993$, $p < .001$), and AlexNet ($\chi2 = 557.143$, $p < .001$).

Table 5: Chi Square Test

Model Comparison	Chi-Square	p-value
ResNet50 vs VGG16	515.1	0.000

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

ResNet50 vs InceptionV3	563	0.000
ResNet50 vs AlexNet	557.1	0.000

The highly significant p-values provide strong evidence that real-world performance ratings were substantially higher for ResNet50 compared to the other CNN models. Examination of crosstabulations showed higher frequencies of 4 and 5 ratings for ResNet50, confirming its superior real-world application based on the questionnaire data. The chi-square tests reinforce ResNet50's advantages in terms of accuracy, efficiency, and applicability.

Linear Regression

A multiple linear regression analysis was conducted to evaluate how well the ResNet50 model characteristics predict image classification accuracy. The independent variables included in the model were training time, convergence rate, model size, generalization error, and real-world performance rating. The dependent variable was accuracy percentage. The results of the regression are presented in Table 6. The model was statistically significant, with $F(5, 380) = 143.011$ and Sig. < .001. The R-squared value was .653, indicating the predictor variables explained 65.3% of the variance in accuracy.

Table 6: Regression Analysis

Dependent Variable: Accuracy%				
Variables	B	Std. Error	T	Sig.
(Constant)	-11.177	4.003	-2.792	0.006
Training Time (hrs)	0.341	0.167	2.040	0.042
Convergence Rate	76.901	21.450	3.585	0.000
Model Size (millions)	-0.054	0.073	-0.741	0.459
Generalization Error (%)	-0.025	0.263	-0.095	0.924
Real-World Performance	16.901	1.079	15.668	0.000
R			0.808	
R2			0.653	
Adjusted R2			0.648	
F			143.011	
Sig-F			0.000	

Among the predictors, real-world performance rating had the strongest positive relationship with accuracy, recording a highly significant regression coefficient ($B = 16.901$, $p < .001$). This suggests models that were rated higher for applicability also achieved greater accuracy. Convergence rate was also a significant positive predictor ($B = 76.901$, $p < .001$), as faster convergence was associated with higher accuracy. Training time had a small but significant positive coefficient were ($B = 0.341$, $p < .05$), implying longer training corresponds to marginal accuracy gains. Meanwhile, model size and generalization error did not have statistically significant effects on accuracy in the regression model.

DISCUSSION

The data analysis findings provide strong evidence that ResNet50 significantly improves image classification accuracy and real-world performance compared to earlier CNN architectures like VGG16, InceptionV3, and AlexNet.

The quantitative results demonstrate ResNet50 achieves superior accuracy on benchmark tests, with around 3-4% higher top-1 classification accuracy on average based on the questionnaire ratings. Additionally, ResNet50 showed faster training times, better convergence rates, lower generalization error, and smaller model sizes in terms of parameters. This aligns with previous research on the advantages of ResNet50's residual learning framework for training deeper networks (He et al., 2016).

The higher questionnaire ratings for real-world usage further verify the benefits of ResNet50 for commercial computer vision applications. The chi-square tests showed significantly higher distributions of ratings for ResNet50 compared to the other models. Qualitative feedback emphasized ResNet50's ability to learn more complex features due to its depth, leading to substantial accuracy gains in tasks like object recognition and facial identification.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Notably, ResNet50 achieved the accuracy and performance improvements while requiring far fewer parameters than other models like VGG16. The regression analysis demonstrated the inverse relationship between number of parameters and accuracy. Thus, the residual architecture does not simply allow for more parameters, but actually improves efficiency. This was further reflected in the superior convergence rates and training times for ResNet50.

Overall, the findings strongly support the research hypothesis that ResNet50 significantly enhances image classification accuracy over prior CNNs. Both the quantitative metrics and qualitative insights highlight the benefits of ResNet50's unique residual design for learning hierarchical representations. The results provide a comprehensive assessment of the impact of using state-of-the-art deep learning models like ResNet50 for computer vision applications.

Conclusion

This research provided compelling evidence that the advanced ResNet50 architecture substantially elevates image classification accuracy and real-world performance compared to earlier CNN models. Quantitative analysis of questionnaire data from 386 experts showed ResNet50 achieves 3-4% higher accuracy on average versus VGG16, InceptionV3, and AlexNet. Statistical tests confirmed these gains are significant across benchmark and applied scenarios. Qualitative feedback highlighted ResNet50's innovative residual learning approach enables training deeper neural networks of 50 layers, allowing more sophisticated feature extraction from images. Critically, ResNet50 attains superior accuracy with fewer parameters than predecessors, demonstrating improved efficiency. The residual connections mitigate degradation problems in very deep networks. Overall, the study quantified significant advantages of ResNet50 over previous CNNs in terms of accuracy, efficiency, and applicability due to its advanced deep residual design. The results will assist researchers and practitioners in selecting optimal deep learning architectures and driving future enhancements.

REFRENCES

[1] Collis, J., & Hussey, R. (2013). Business Research: A practical guide for undergraduate and postgraduate students. Palgrave Macmillan.
[2] Hair Jr, J. F., Wolfinbarger, M., Money, A. H., Samouel, P., & Page, M. J. (2015). Essentials of business research methods. Routledge.
[3] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
[4] Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International conference on machine learning (pp. 448-456).
[5] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.
[6] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.
[7] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Berg, A. C. (2015). Imagenet large scale visual recognition challenge. International journal of computer vision, 115(3), 211-252.
[8] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations.
[9] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).
[10] Bryant, F. B., & Satorra, A. (2012). Principles and practice of scaled difference chi-square testing. Structural equation modeling: A multidisciplinary journal, 19(3), 372-398.