Deep-Learning Memory Systems: Rigorously Searching for General Principles Behind Skip Connections

By Vincent Abbott

1. Introduction

Deep learning has enabled significant progress in a wide range of applications, from image and speech recognition to natural language processing and autonomous vehicles. However, model design choices can make or break an implementation. For example, models must be powerful enough to incorporate and manipulate data in helpful ways yet frugal enough to avoid overfitting (learning the noise in the data rather than the underlying pattern) and reduce training times.

One technique of particular interest are skip connections, which have demonstrated significant practical utility in various deep-learning architectures. Skip connections allow information from early layers in a network to bypass some intermediate layers and be combined with later ones. This approach has been instrumental in innovations such as highway networks (1), residual networks (3), U-Nets (2), and densely connected convolutional networks (4). Residual networks, for example, are deep learning models that use skip connections to alleviate the vanishing gradient problem, which hampers the training of very deep networks. U-Nets are a type of convolutional neural network designed for image segmentation tasks, utilizing skip connections to combine information from different scales of an image.

The primary goal of this research is to identify general principles behind skip connections that can inform better deep-learning model design choices. To achieve this, we will survey the literature on skip connections, extracting key insights and testing them through experiments. By the end of this research, we aim to present a set of experimentally validated design principles that can guide future machine learning work.

Model power, or the ability to implement arbitrary functions, relies heavily on the number of available parameters and available training time. Powerful models, however, may suffer from overfitting and require more computational resources, leading to environmental and financial costs (9). In contrast, model economics focuses on the efficient use of available parameters and training time, emphasizing the need for stable and focused models. As a result, techniques such as batch normalization (7), which controls the mean and variance of inputs to activation functions, have become essential for state-of-the-art performance by stabilizing activations and allowing for more stable training.

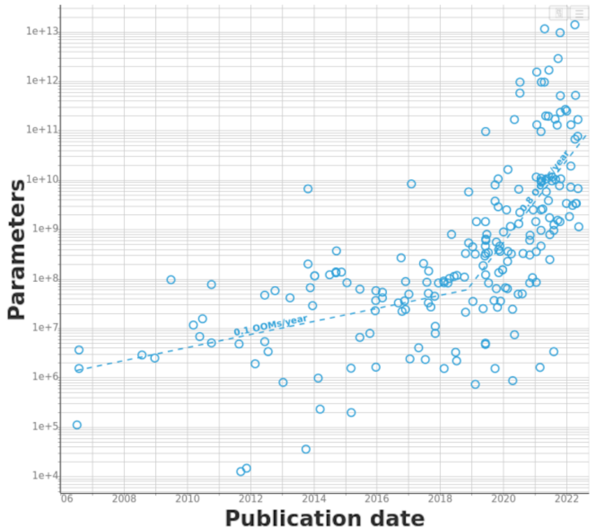

Figure 1: The number of parameters in machine learning models has increased exponentially over time (12). The return on proportional efficiency gains for a given model size grows exponentially yearly.

Among the various contributions to model economics, skip connections stand out due to their ability to focus available parameters on finding new patterns while preserving old ones. Moreover, skip connections provide direct access to patterns extracted early in the network, reducing parameter sensitivity and stabilizing training. Investigating skip connections rigorously to extract general principles that can enhance the quality of machine learning models is, therefore, a valuable avenue to explore.

2. The Research Problem

The key research question we aim to address is: What makes skip connections so effective in improving deep learning model performance and training stability without requiring increased data or processing power? Addressing this question could help tackle the real-world problem of balancing model power with model economics, enabling more efficient and accessible deep-learning applications. This research aims to identify general principles which inform why and when skip connections should be used. The focus on finding principles is motivated by the seemingly unintuitive nature of residual networks. Despite machine learning being around for decades, this computationally effective method is less than eight years old. We will survey models that have made notable contributions to model economics, extract potential principles, formulate hypotheses, and test these hypotheses through experiments. By the end of this process, we aim to establish principles that can guide the development of more efficient and effective deep-learning models.

2.1. Real World Significance

Improving model economics in machine learning is crucial for efficiently scaling models and addressing complex problems, such as artifact-free video generation, which involves creating high-quality, seamless video content. Better model economics can make the training of models for such tasks more tractable and enable more strategic use of available parameters. Additionally, enhancing performance without relying on increased access to clean data and computing resources can reduce the financial and environmental costs (9) associated with machine learning research and commercial applications. Lower financial costs can increase accessibility for new researchers and allow for more experimental iterations while reducing environmental concerns related to high power consumption and emissions from training and querying large models.

Skip connections are particularly relevant in the real-world context due to their significant contribution to model economics. For example, before the introduction of highway networks (1), models struggled to benefit from increased depth. Residual networks (3) further improved on this concept by using skip connections without requiring additional parameters, leading to a strict improvement in performance. In the following sections, we will review skip connections and their contributions to model economics and discuss this principle-focused approach's specific advantages and real-world impact.

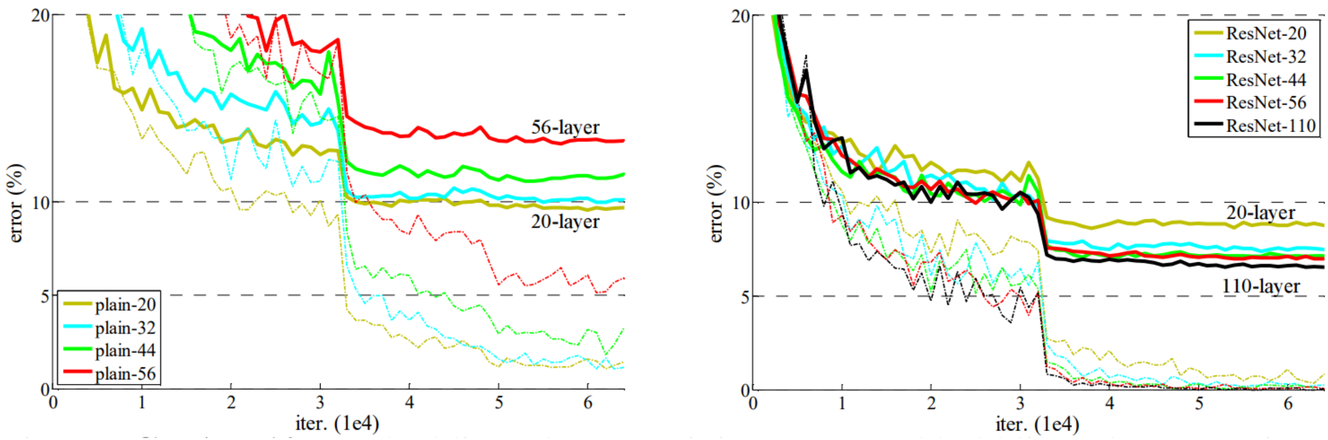

Figure 2: From the ResNet paper, the performance advantage of residual connections can be clearly seen. On the **left** are plain networks, whose training success suffers at increased depth. On the **right**, networks with residual skip connections strictly benefitted from depth. Models were tested on CIFAR-10. Dashed lines denote training error, and bold denote test error.

2.2. Related Work

Skip connections, which preserve information for later reintroduction, play a critical role in model economics. Several notable models have used this technique to improve model performance without relying on increased size. For example, U-Nets (2), introduced in May 2015 for image segmentation, feature a "U" shape, down-sampling (reducing resolution) and then up-sampling (increasing resolution) images while reintroducing earlier layer information. Initially designed for biomedical image processing, U-Nets have become essential for image segmentation and processing tasks, including modern image generation systems like Dall-E, Midjourney, and Stable Diffusion (5,6).

Highway networks, introduced in November 2015, combine the input and output of each layer in a ratio derived from the input and learned parameters. Shortly after, residual networks were introduced in December 2015, with the original paper receiving nearly 160,000 citations. These networks calculate additive perturbations to the input rather than complete transformations, allowing models to benefit from depth and maintain information over specific dimensions.

The emphasis on dates highlights the recentness of these innovations, each with tens of thousands of citations. In 2015, models first exceeded depths of 100 layers. Skip connections have since inspired further work, such as densely connected convolutional networks for image processing (4) (January 2018), which grant each layer access to all output features from previous layers. This approach improves training stability and performance with fewer parameters. Transformers (8), which revolutionized natural language processing, also employ a form of skip connections, encoding input strings and feeding them into each decoder layer.

Various papers have proposed reasons for the effectiveness of skip connections, such as feature reuse, availability of identity-like layers, and enabling selective information flows (1,3,4). Commonly agreed aspects include allowing parameters to focus on new patterns and stabilizing training with connections insensitive to parameters. This research aims to assess the reasons suggested by these works, propose and test principles, and investigate which justifications hold more validity than others.

2.3. Advantages of this Approach

This research will focus on finding general principles, a specific approach with additional real-world benefits. Skip connections, despite their simplicity and utility, are novel. U-Nets, highway networks and residual networks were all introduced in 2015, after AlexNet (10), GANs (11) or the decades of research struggling to train networks without modern processors. This suggests they are in some way unintuitive. If a set of design principles fails to make the intuition behind good design choices clear, those principles should be updated. This shortfall motivates us to search for general principles which can inform better design choices. The rigorous hypothesizing and testing we will use make generalization more likely and the work more informative, increasing the likelihood of aiding model economics in future work.

The general principles derived from this research can directly impact the development of future models by guiding researchers in making informed design decisions that can lead to improved model performance and training efficiency. Moreover, clear design principles lower the barrier to entry for people interested in the field. Without such principles, teaching and learning can be ad-hoc and accessible only to the most motivated students willing to search for patterns themselves. An established framework that can intuitively explain residual networks and other architectures empowers newcomers to more easily understand the rapid development of new models. Additionally, these principles can foster interdisciplinary collaboration, as clearer design principles can be more easily understood by researchers from different fields, leading to innovative solutions that combine various domains of expertise.

3. Methodology

This research focuses on discovering design principles for the appropriate use of skip connections in machine learning models by employing a three-step methodology: surveying the literature, proposing principles, and evaluating principles through experiments. The literature review provides a foundation for understanding the context and reasoning behind successful skip connection implementations, while the proposed principles synthesize insights and offer testable hypotheses. Finally, these principles are evaluated by conducting experiments and refining and validating the principles based on the results. Ultimately, we aim to create a comprehensive framework to guide the effective use of skip connections in model economics. This research aims to contribute to human knowledge by enhancing the understanding of skip connections, making their design more accessible and intuitive for newcomers and experts in the machine learning field.

3.1. Surveying the Literature

The first part of the methodology focuses on surveying the literature to identify papers that have made significant contributions to model economics, particularly those involving skip connections. This comprehensive literature review will provide a strong foundation for understanding the current state of knowledge and help derive design principles for skip connections. The following steps outline the process of surveying the literature:

1. **Selection Criteria:** Establish criteria for selecting papers that are relevant, have made immense contributions to model economics, and involve the use of skip connections. Ensure the chosen papers include those with innovative modifications which are still being used. Also include papers that have significantly contributed to model economics, even if they do not explicitly use skip connections.
2. **Identification of Relevant Papers:** Conduct a thorough search for papers in academic databases, conference proceedings, and other reputable sources. Use keywords related to skip connections, model economics, and machine learning architectures to find relevant papers.
3. **Review and Analysis of Papers:** Read the identified papers and critically analyze their content. Focus on understanding the reasoning behind the improved performance and how skip connections were utilized in each paper.
4. **Extraction of Key Information:** Record essential information from each paper, including the reasoning for improved performance, the context of skip connections, the model innovations used, and the results achieved. This information will help derive design principles and compare different approaches.
5. **Classification and Organization:** Organize the collected information based on common explanations, methodologies, and skip connection types. This organization will facilitate the comparison of different approaches and help in identifying patterns and trends in the literature.
6. **Synthesis of Findings:** Combine the findings from the literature review to develop an understanding of the current state of knowledge in the field. Identify gaps in knowledge, areas of consensus, and potential avenues for further exploration.

By following these steps, the literature survey will provide a comprehensive understanding of the existing knowledge on skip connections and their role in model economics. This understanding will serve as the foundation for proposing and evaluating design principles in the subsequent parts of the methodology.

3.2. Proposing Principles

The second part of the methodology focuses on proposing design principles for skip connections based on the findings from the literature review. These principles will be used to guide the appropriate use of skip connections and provide intuitive explanations for their advantages. The following steps outline the process of proposing principles:

1. **Identification of Reasoning:** Review the collected information from the literature survey to identify the various reasonings that authors have provided for the improved performance of their models using skip connections. Take note of both explicit and implicit reasoning and any conjectures made by the authors.
2. **Formulation of Potential Principles:** Using the identified reasoning from the literature, formulate a set of potential principles that could explain the success of skip connections. Ensure that these principles are clear, concise, and have intuitive explanations.
3. **Integration of Novel Observations:** Besides the principles derived from the literature, consider any novel observations or patterns identified during the literature review. Use these insights to propose additional principles that could guide the use of skip connections.
4. **Evaluation of Principles:** Assess the proposed principles based on their ability to explain the success of skip connections in the literature. Consider their generality, intuitiveness, and compatibility with the existing knowledge in the field.
5. **Testability of Principles:** Ensure that each proposed principle offers testable hypotheses that we can experimentally evaluate in the next part of the methodology. This will allow for validating or rejecting the principles based on empirical evidence.
6. **Refinement and Consolidation:** Refine and consolidate the proposed principles by eliminating redundancies, inconsistencies, or principles that lack intuitive explanations. The resulting principles should be coherent, comprehensive, and representative of the knowledge gained from the literature review.

By following these steps, the proposed principles will provide a solid foundation for understanding the appropriate use of skip connections and their advantages. These principles will be further evaluated and validated through experimentation in the third part of the methodology, ensuring their applicability and relevance to the field of machine learning.

3.3. Evaluating Principles through Experiments

The third part of the methodology aims to evaluate the proposed principles by conducting experiments that test the hypotheses derived from these principles. This process will provide empirical evidence supporting or refuting the principles, ensuring their validity and applicability in the context of skip connections. The steps involved in this part of the methodology include the following:

1. **Hypothesis Generation:** For each proposed principle, generate testable hypotheses that predict the performance or training speed of various model designs involving skip connections. These hypotheses should be specific, measurable, and directly related to the principles being evaluated.
2. **Experimental Design:** Develop an experimental design that allows for the systematic evaluation of the hypotheses. This includes selecting the appropriate models, architectures, datasets, and performance metrics to be used in the experiments. Ensure that the experimental design is controlled, reproducible, and capable of providing robust evidence for or against the hypotheses.
3. **Experiment Execution:** Conduct the experiments as per the designed plan, carefully documenting the setup, parameters, and results for each experiment. Ensure that the experiments are run consistently and accurately, maintaining high rigor and control throughout the process.
4. **Data Analysis:** Analyze the experimental results, comparing the performance and training speed of the different model designs in relation to the hypotheses being tested. Use appropriate statistical methods to determine the significance of the observed differences and draw conclusions about the validity of the hypotheses.
5. **Principle Validation:** Validate or reject the proposed principles based on the experimental results and data analysis. Successful principles should have hypotheses supported by experimental evidence, while unsuccessful principles should be refined or discarded.
6. **Refinement and Synthesis:** Refine the remaining successful principles by further analyzing the experimental results and identifying common themes or patterns. Synthesize the principles into a cohesive and comprehensive framework that captures the essential insights about skip connections and their advantages.

Through this experimental evaluation process, the proposed principles will be thoroughly tested and validated, ensuring their relevance and applicability in guiding the use of skip connections in machine learning models. The final set of principles will provide a robust and intuitive understanding of skip connections, contributing to the improvement of model economics and the broader field of machine learning.

4. Evaluation Criteria

The primary goal of this research is to rigorously identify and develop design principles that offer clear intuition for the appropriate use of skip connections in machine learning models. To evaluate the success of the research, we will consider the following criteria:

1. **Validity of Principles:** The derived design principles should demonstrate success across various experiments that test their validity and compare them against alternative

explanations. The experimental results should provide evidence in support of the proposed principles and indicate their relevance in different contexts.
2. **Accessibility of Intuition:** The principles should be easy to understand and digest, making machine learning design more accessible to newcomers. The explanations underpinning these principles should be straightforward, enabling learners to grasp the key ideas before delving deeper into the literature.
3. **Generalizability:** The principles should not be ad-hoc, meaning they should explain a wide range of situations and remain relevant in various applications. This criterion emphasizes the importance of a robust methodology that considers multiple explanations, establishes principles and hypotheses prior to testing, and avoids overfitting to specific experimental results.
4. **Guidance on Inappropriate Use:** The principles should also inform us when skip connections are not helpful or recommended. A comprehensive set of principles should include both positive cases (when skip connections should be used) and negative cases (when they should not be used). This distinction ensures that the principles provide actionable guidance for model design.
5. **Unified Explanation:** If multiple principles are proposed, they should be connected by a unified explanation or intuition. This cohesion will help ensure that the principles form a coherent framework, making it easier for practitioners to apply them in their work.

By offering improved design principles for the appropriate use of skip connections in machine learning models, this project will contribute to human knowledge in several significant ways. Satisfying the evaluation criteria will give confidence that a meaningful contribution has been made by addressing the following aspects:

1. **Enhancing Model Design and Performance:** Developing a robust set of design principles will enable researchers and practitioners to optimize machine learning models more effectively. By understanding when and how to use skip connections appropriately, they can create models with improved performance, generalization capabilities, and training efficiency.
2. **Facilitating Knowledge Transfer:** Clear and accessible principles will make it easier for newcomers and experts to grasp the core concepts related to skip connections. This understanding will promote the transfer of knowledge within the field, accelerating innovation and the adoption of best practices in machine learning model design.
3. **Encouraging Rigorous Research:** This project will promote a more systematic and thorough approach to machine learning research by emphasizing a rigorous methodology and evaluation criteria. This emphasis will contribute to developing a stronger scientific foundation within the field and encourage researchers to seek generalizable and reliable explanations.
4. **Bridging Theory and Practice:** Developing design principles that are both theoretically sound and practically relevant will strengthen the connection between machine learning theory and real-world applications. This bridge will help ensure that theoretical advancements translate into tangible improvements in the performance of

machine learning models, thereby maximizing their impact across various domains.

5. **Inspiring Future Research:** The principles and insights derived from this project can inspire new avenues of research in machine learning, leading to further advancements in model design, optimization techniques, and understanding of the underlying mechanisms that drive successful learning. In addition, by contributing novel knowledge, this project will stimulate the intellectual curiosity of researchers and promote continued exploration within the field.

By satisfying the evaluation criteria and addressing these aspects, the project will make a meaningful contribution to human knowledge. Furthermore, the improved design principles will enhance our understanding of skip connections and help drive progress in machine learning, ultimately benefiting numerous applications across various domains.

Acknowledgements: ChatGPT was an invaluable assistant in preparing this work, and refining drafts into final versions.

5. Bibliography

1. Srivastava RK, Greff K, Schmidhuber J. Highway Networks. CoRR. 2015;abs/1505.00387.
2. Ronneberger O, P.Fischer, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer; 2015. p. 234–41. (LNCS; vol. 9351).
3. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. CoRR. 2015;abs/1512.03385.
4. Huang G, Liu Z, Weinberger KQ. Densely Connected Convolutional Networks. CoRR. 2016;abs/1608.06993.
5. Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B. High-Resolution Image Synthesis with Latent Diffusion Models. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2022.
6. Ho J, Jain A, Abbeel P. Denoising Diffusion Probabilistic Models. CoRR. 2020;abs/2006.11239.
7. Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. CoRR. 2015;abs/1502.03167.
8. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention Is All You Need. CoRR. 2017;abs/1706.03762.
9. Patterson DA, Gonzalez J, Le QV, Liang C, Munguia L-M, Rothchild D, et al. Carbon Emissions and Large Neural Network Training. CoRR. 2021;abs/2104.10350.
10. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. In: Pereira F, Burges CJ, Bottou L, Weinberger KQ, editors. Advances in Neural Information Processing Systems. Curran Associates, Inc.; 2012.
11. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative Adversarial Networks. 2014.

12. Villalobos P, Sevilla J, Besiroglu T, Heim L, Ho A, Hobbhahn M. Machine Learning Model Sizes and the Parameter Gap. 2022.