DOI: 10.6977/IJoSI.202506_9(3).0001
M. Cosgun, K. Altun/Int. J. Systematic Innovation, 9(3), 1-7 (2025)

Innovative solutions for convolutional neural network performance: A TRIZ-based reverse engineering approach

Merve Cosgun, Koray Altun*

Department of Industrial Engineering, Bursa Technical University, Bursa, Turkey

*Corresponding author E-mail: koray.altun@btu.edu.tr

(Received 23 September 2024; Final version received 31 December 2024; Accepted 14 January 2025)

Abstract

Convolutional neural networks (CNNs) are widely used in computer vision for tasks like image classification and detection. These models work well when the number of image classes is small, but as the number of classes increases, accuracy tends to drop due to overfitting. There are several methods to address this issue, such as data augmentation, preprocessing, class weighting, transfer learning, and adjusting technical parameters. This study introduces a novel approach utilizing the theory of inventive problem-solving (TRIZ) methodology to systematically analyze and enhance these existing methods. Using reverse engineering, we deconstructed current solutions and aligned them with TRIZ principles to propose more innovative and effective approaches for improving CNN performance. The results show that TRIZ provides a structured and creative framework for solving accuracy decline issues in CNN models, offering the potential for broader applications in other machine learning architectures.

Keywords: Convolutional Neural Network, Image Classification, Reverse Engineering, Theory of Inventive Problem Solving

1. Introduction

Convolutional neural networks (CNNs) are widely used in computer vision tasks like image classification due to their high performance in learning patterns from data (LeCun et al., 2015). However, as the number of classes in a dataset increases, model accuracy tends to decrease. This drop in accuracy is often caused by overfitting, where the model becomes too specialized in the training data and struggles to generalize to new, unseen data (Krizhevsky et al., 2017). Solving this problem is critical for applications that require accurate classification across many classes, such as medical diagnosis, autonomous driving, and facial recognition systems (Guo et al., 2019; Litjens et al., 2017).

Several approaches have been proposed to address this issue, including data augmentation, preprocessing techniques, and transfer learning (Shorten & Khoshgoftaar, 2019).

"Data augmentation," for example, increases the diversity of the training data by applying transformations such as rotation, scaling, and flipping to existing images. This approach has proven effective in applications such as medical imaging, where obtaining large datasets is difficult (Perez & Wang, 2017).

Traditional augmentation techniques, such as rotation, flipping, and cropping, have been widely used in image classification and segmentation tasks. However, more innovative strategies are continuously being developed to enhance augmentation effectiveness. For instance, Alomar et al. (2023) introduced a new "random local rotation" technique, which improves data diversity while minimizing the boundary artifacts commonly caused by traditional rotation methods. These advancements in augmentation help CNNs generalize better, especially in tasks with limited data availability.

Several approaches have been proposed to improve data augmentation, particularly with automated techniques. Automated data augmentation (AutoDA) methods have been increasingly studied, as they can automatically discover optimal augmentation strategies tailored to specific datasets. For example, a recent comprehensive survey by Yang et al. (2023)

categorizes existing AutoDA methods, highlighting their efficiency in improving image classification tasks by reducing manual intervention and increasing model performance through learned augmentation policies. These AutoDA methods present a promising direction for enhancing data diversity and model generalization while reducing human error in the augmentation process.

"Preprocessing" techniques, such as normalization and cropping, are used to refine input data before training, ensuring that the model receives consistent and high-quality information (Kamnitsas et al., 2017).

Image preprocessing is essential for improving CNN performance by reducing noise and enhancing data quality. Techniques such as noise reduction, histogram equalization, and image hashing have shown notable accuracy improvements in facial recognition tasks, with gains of over 4% in some cases (Tribuana et al., 2024).

"Transfer learning," on the other hand, leverages knowledge from pre-trained models to enhance performance in specific tasks, particularly when labeled data is limited, allowing models pre-trained on large datasets to be fine-tuned for specific tasks (Atasever et al., 2023; Pan & Yang, 2010; Tan et al., 2018).

More innovative approaches are needed to tackle the root of the problem in a systematic way. This is where the theory of inventive problem-solving (TRIZ) methodology comes into play. Developed in engineering, TRIZ offers a structured approach to identifying contradictions and proposing creative solutions based on inventive principles. The method has proven effective in solving technical problems across various fields, yet its application in machine learning, particularly for improving CNN performance, remains underexplored.

The aim of this study is to address the issue of decreasing accuracy in CNN models as the number of classes increases. By applying the TRIZ methodology, we aim to develop innovative and systematic solutions to improve CNN performance. Through reverse engineering, we analyzed current techniques, such as data augmentation, preprocessing, class weighting, and transfer learning, and aligned these with TRIZ principles to propose more effective solutions.

Recent advancements in artificial intelligence (AI) have increasingly focused on overcoming the challenges posed by small datasets, which are prevalent in fields where data collection is restricted. The study by Brad and Brad (2023) explores the use of TRIZ methodologies to address this issue, offering inventive strategies for enhancing AI model performance under these constraints. While their research underscores the adaptability of TRIZ in optimizing AI with limited data, our study extends the application of TRIZ to tackle accuracy degradation in CNN models as class counts increase. This divergence highlights the broad applicability of TRIZ principles across different AI challenges.

The contribution of this study lies in the novel application of TRIZ methodology to the field of machine learning, specifically in solving the overfitting problem in CNNs with increasing class counts. TRIZ, traditionally applied in engineering, provides a structured approach to identifying contradictions and generating innovative solutions, which has not been widely explored in CNN performance issues.

The remainder of this paper is organized as follows: Section 2 details the methodology; Section 3 presents the implementation; Section 4 provides the discussion; and Section 5 concludes the paper, summarizing the findings and offering directions for future research.

2. Methodology

The TRIZ methodology provides a systematic approach to innovation that involves analyzing and categorizing thousands of patents to uncover universal principles of invention. Central to TRIZ is its distinctive method for tackling technical problems by converting specific situations into broader, conceptual challenges. This process requires breaking down the problem into its core elements and then applying TRIZ's set of inventive principles and proven solutions to devise a conceptual solution (Gadd, 2011). The TRIZ methodology is traditionally outlined as:

(i) Specific problem: The initial stage, is where the specific technical problem is identified.
(ii) Conceptual problem (39 parameters): The problem is generalized to a conceptual level by identifying relevant engineering parameters.
(iii) Conceptual solution (40 principles): Solutions are developed based on the 40 inventive principles of TRIZ.
(iv) Specific solution: The conceptual solution is then translated back into a specific practical solution for the initial problem.

The process starts by breaking down a real-world problem into a conceptual format. This simplification allows for aligning the problem with TRIZ's effective solutions, which rely on structured principles instead of random brainstorming ideas. After finding a conceptual solution, it is then polished and converted into a practical solution that specifically addresses the initial problem.

The TRIZ methodology focuses on transforming specific, real-world problems into conceptual challenges, which can then be matched with systematic solutions. Sheu & Lee (2011) proposed a structured process for innovation that incorporates TRIZ principles to help break down complex problems and develop creative solutions. Their work highlights the importance of using TRIZ tools, such as the contradiction matrix and inventive principles, to ensure that the problem-solving process is both organized and comprehensive. By following their systematic process, innovators can consistently arrive at effective solutions for technical challenges.

In this study, we approach TRIZ from a "reverse engineering" perspective, applying its principles not just to generate new solutions but also to reinterpret and reanalyze existing solutions found in the literature. By doing so, we aim to provide a more comprehensive framework for problem-solving that bridges past solutions with inventive methodologies. The methodology proposed in this study is as follows;

(i) Specific solution: Start from an existing solution or product.
(ii) Conceptual solution analysis (40 principles): Deconstruct the solution to understand how TRIZ principles are or can be applied.
(iii) Identification of contradictions (39 parameters): Identify any existing or potential contradictions that the current solution might be causing or not addressing.
(iv) Revised problem statement: Define or redefine problems based on insights gained from the analysis and contradiction identification.

Reverse engineering within the TRIZ framework involves deconstructing existing technical solutions to understand their core principles and then matching these with TRIZ's 40 inventive principles and the contradiction matrix. This approach allows us to assess how well these existing solutions align with TRIZ's systematic process and identify opportunities for improvement or further development. For instance, a solution that addresses one specific technical contradiction may have untapped potential for solving additional contradictions when viewed through the lens of TRIZ.

Fig. 1 demonstrates how the traditional TRIZ process is a forward-thinking approach, starting from problem identification and moving toward a solution. In contrast, the reverse engineering process begins with an existing solution, analyzing it through the TRIZ lens to uncover deeper insights and potentially redefine the problem or improve the solution.

3. Implementation

In this section, we take a different path from the usual forward-thinking problem-solving approaches. This is where reverse engineering comes into play. Reverse engineering, in essence, involves working backward to deconstruct an existing solution to understand its foundational principles. In doing so, we can uncover hidden opportunities for improvement or discover alternative solutions that may not have been obvious initially. Instead of building a solution from scratch, we analyze what already exists, break it down, and explore how it aligns with TRIZ's 40 inventive principles and contradiction matrix.

By reversing the usual flow of thought, we can gain deeper insights into how existing solutions operate and how they can be enhanced or adapted for broader applications. This method is particularly useful in complex problems, such as improving the accuracy of CNN models, where conventional methods may overlook indirect contradictions or potential improvements that TRIZ can highlight.

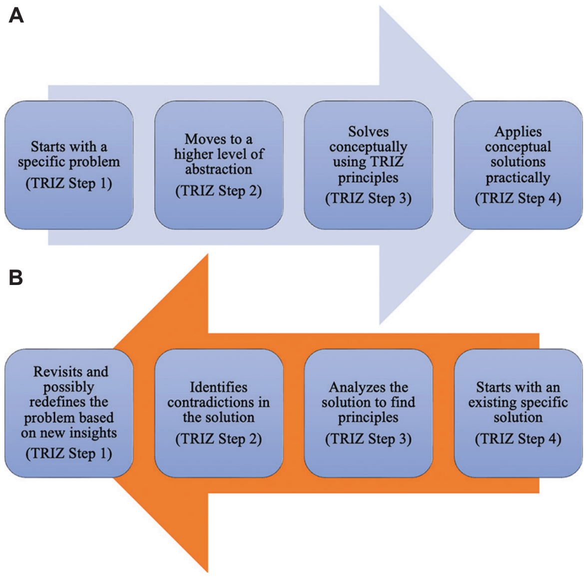

Fig. 1. A comparison of the flows of a (A) TRIZ process and a (B) TRIZ-based reverse engineering process
Abbreviation: TRIZ: Theory of inventive problem solving.

3.1. TRIZ Step 4: Identifying Specific Solutions

In TRIZ step 4, we focus on identifying concrete, practical solutions to address the issue of decreasing accuracy in CNN models as the number of image classes increases. From a reverse engineering perspective, we analyze how existing methods have been applied to similar problems and how TRIZ principles can guide the improvement of these methods.

One effective solution is "data augmentation," which involves expanding the dataset by generating new training examples through techniques such as rotating, flipping, scaling, cropping, and adding noise. We can see that "Principle 20: Continuity of Useful Action" fits well here, as the method continuously provides useful variations of the data that help the model learn better. “Principle 15: Dynamics” also applies because the transformations increase the adaptability of the model to diverse input scenarios.

Similarly, “dimensional adjustment” is another common preprocessing technique that improves consistency across classes by resizing or normalizing images. This approach aligns with “Principle 17: Another Dimension,” which suggests modifying or using different dimensions to solve a problem. By ensuring that images are of uniform size, we reduce the variability in the input data, which enhances the model’s ability to make accurate predictions.

In cases where there is a class imbalance, “class weighting” can improve accuracy by giving more importance to underrepresented classes. This method reflects “Principle 35: Parameter Changes,” which involves adjusting key parameters to achieve the desired result. Assigning weights to classes based on their representation helps the model treat all classes fairly, reducing bias and improving performance.

Finally, “transfer learning” offers a powerful way to reuse pre-trained models on new tasks, particularly when data is scarce. This approach aligns with “Principle 24: Intermediary,” which suggests using an intermediary to assist with problem-solving. The intermediary model speeds up learning and improves performance, especially when the new task shares similarities with the original one.

3.2. TRIZ Step 3: Identifying Conceptual Solutions

In TRIZ step 3, the goal is to select a conceptual solution based on TRIZ principles that fits the problem. Using a reverse engineering approach, we look back at how similar problems have been solved in the past and apply those insights to find conceptual solutions for the current issue of CNN accuracy decline. Table 1 lists all the candidate principles. Table 2 shows the candidate principles that match the specific solutions presented in step 4.

The candidate principles listed in Table 1 were selected through a structured TRIZ-based process. Initially, principles annotated with the letter “a” (i.e., 15, 17, 20, 24, and 35) were identified based on their relevance to the solution alternatives in Table 2. These principles were mapped to engineering characteristics using a contradiction matrix, highlighting key pairs (Table 3).

For “data augmentation,” we selected Principles 15 and 20 because these principles describe how system flexibility and continuous beneficial actions can enhance performance.

When considering “dimensional adjustment,” Principle 17 stood out as the most appropriate. Adjusting the dimensions of input data can help reduce the computational complexity and make the data more uniform, thereby improving the model’s ability to classify images.

In the case of “class weighting,” Principle 35 was chosen because it directly addresses the problem of imbalanced datasets. Modifying the class weights allows the model to handle rare classes more effectively, which is essential for improving overall accuracy.

Lastly, for “transfer learning,” Principle 24 is highly relevant. Transfer learning enables models to leverage previously learned knowledge, significantly reducing training time and improving accuracy.

Table 1. Candidate principles

No.	Principle	Definition
1	2	Taking out
2	3	Local quality
3	13	The other way round
4	15[a]	Dynamics
5	17[a]	Another dimension
6	18	Mechanical vibration
7	20[a]	Continuity of useful action
8	23	Feedback
9	24[a]	Intermediary
10	27	Cheap short-living objects
11	28	Mechanic substitution
12	29	Pneumatics and hydraulics
13	30	Flexible shells and thin films
14	33	Homogeneity
15	35[a]	Parameter changes
16	36	Phase transitions
17	37	Thermal expansion

Note: [a]Candidate principles.

Table 2. Candidate principles and solution matching

No.	Solution alternatives	TRIZ candidate principle
1	Data augmentation	15, 20
2	Preprocessing	17
3	Assigning class weights	35
4	Transfer learning	24

Abbreviation: TRIZ: Theory of inventive problem-solving.

3.3. TRIZ Step 2: Defining the Conceptual Problem

In TRIZ, step 2 involves identifying the engineering characteristics that are in conflict, resulting in a technical contradiction. Engineering

Table 3. Candidate pairs

Candidates	**28**	**29**	**35**	**39**
10	10, 23, 24[a], 35[a]	28, 29, 36, 37	15a, 17[a], 18, 20[a]	3, 28, 35[a], 37
26	2, 13, 28	30, 33	3, 15a, 29	3, 13, 27, 29

Note: [a]Candidate principles.

characteristics refer to specific technical parameters or features of a system. A contradiction occurs when improving one characteristic negatively impacts another. In our case, increasing the number of image classes in a CNN model may lead to a decrease in accuracy—these are two conflicting engineering characteristics.

Based on the guidance from the candidate principles identified in TRIZ step 3, we selected a pair of engineering characteristics. Table 4 outlines the engineering characteristics obtained from the normal analysis of the problem. Two major contradictions were identified: one representing the increase in the number of classes (related to engineering characteristic 26, "quantity of substance") and another representing the decrease in model accuracy (related to engineering characteristic 35, "adaptability or versatility"). These contradictions must be addressed to improve the model's performance.

Table 3 shows the conceptual solution combinations derived from these contradictions. The contradiction pair 26 and 35, guided by Principles 15, 17, and 20, was determined to be the most relevant for this study. Engineering characteristic 26 corresponds to the challenge of managing an increasing number of classes while engineering characteristic 35 relates to the issue of decreasing accuracy. These contradiction reflects the balance we seek between increasing model capacity and maintaining high accuracy.

It should be highlighted here that, unlike the traditional TRIZ methodology that starts with engineering contradictions, our approach began with identifying the inventive principles, as shown in Table 3. For instance, the contradiction pair (10, 35) was highlighted due to its strong alignment with Principles 15, 17, and 20.

The inclusion of "Force" was guided by its connection to critical principles such as 15 (Dynamics). While characteristics like "Strength" were considered, no directly related principles were identified, which justified its exclusion.

By identifying the engineering characteristics and their associated contradictions, we can apply TRIZ principles to systematically resolve these conflicts. For example, "Principle 15: Dynamics" helps address flexibility in handling different classes, while "Principle 17: Another Dimension" suggests altering how data are processed to maintain accuracy despite increased complexity.

Table 4. Candidate contradictions

No.	**Eng. char.**	**Definition**
1	10	Force (intensity)
2	26	Quantity of substance
3	28	Measurement accuracy
4	29	Manufacturing precision
5	35	Adaptability or versatility
6	39	Productivity

Abbreviation: Eng. Char.: Engineering characteristic.

In comparison, Brad and Brad (2023) explored contradictions between data quantity and system performance (contradiction pairs 26 & 28) in their study. Their analysis of contradictions between system complexity and performance (contradiction pair 36 & 28) further illustrates how TRIZ can be used to systematically resolve technical challenges by focusing on the core engineering characteristics at play.

4. Discussion

In this study, we applied a TRIZ-based reverse engineering approach to address the problem of decreasing accuracy in CNN models as the number of image classes increases. This methodology enabled a systematic examination of existing solutions, revealing opportunities for innovation by aligning these solutions with TRIZ principles.

We further discuss the insights gained through this approach, with a particular focus on the application of transfer learning and its reinterpretation within the TRIZ framework.

4.1. General Insights

The TRIZ-based reverse engineering shifts the focus from conventional problem-solving to a structured analysis of existing solutions. This approach provides a systematic way to identify and resolve contradictions inherent in CNN models, such as the trade-off between model complexity and accuracy. By deconstructing existing methods, such as data augmentation, preprocessing, and class weighting, we identified their alignment with specific TRIZ principles and proposed refinements that address overlooked challenges.

For instance, data augmentation aligns with "Principle 20: Continuity of Useful Action," as it generates continuous variations in the training data, enabling better model generalization. Similarly, preprocessing techniques, such as dimensional adjustments, align with "Principle 17: Another Dimension," addressing variability in input data to improve consistency and prediction accuracy. These connections demonstrate how TRIZ principles provide a creative and structured framework to optimize existing solutions.

4.2. Transfer Learning and Intermediary Principle

We acknowledge that transfer learning is a widely recognized method for improving CNN performance, particularly in cases with limited labeled data. However, its reinterpretation through TRIZ's "Principle 24: Intermediary" offers new perspectives and applications.

Transfer learning is typically seen as a way to reuse pre-trained models for specific tasks. Within the TRIZ framework, we redefine it as an intermediary that bridges two conflicting needs: (i) the scarcity of labeled data in new tasks, and (ii) the requirement for high accuracy in performance. This reinterpretation positions transfer learning not just as a static tool but also as a dynamic mediator that facilitates the resolution of these contradictions. By emphasizing its role as an intermediary, TRIZ provides a structured perspective for enhancing the applicability of transfer learning.

Viewing transfer learning through the lens of TRIZ enables a broader application of this technique. For example, TRIZ principles encourage creative extensions of pre-trained models to address additional challenges, such as:

(i) Reducing bias in class imbalance: By systematically reweighting features learned by the intermediary model, we can mitigate biases present in underrepresented classes.
(ii) Overcoming noise in preprocessing: Transfer learning can serve as a filter to preprocess noisy data more effectively, guided by TRIZ principles, such as "Principle 35: Parameter Changes."

These reinterpretations highlight how TRIZ inspires creative problem-solving by encouraging researchers to think beyond the conventional uses of established techniques.

5. Concluding Remarks

The TRIZ-based reverse engineering approach, proposed in this study, offers a structured framework for analyzing existing solutions, uncovering contradictions, and proposing innovative refinements. Unlike traditional forward-thinking methods, reverse engineering starts with what already exists, systematically deconstructs these solutions, and applies TRIZ principles to identify untapped opportunities. This methodology provides two major advantages:

(i) Systematic problem analysis: Traditional problem-solving approaches often focus on developing new solutions from scratch. In contrast, reverse engineering allows us to examine existing solutions critically, revealing their underlying contradictions or limitations. By aligning these with TRIZ's 40 inventive principles and contradiction matrix, we can systematically identify areas for improvement or innovation.
(ii) Maximizing existing knowledge: This approach avoids reinventing the wheel by leveraging what is already available. It enables researchers to reinterpret existing solutions through the lens of TRIZ principles, uncovering new opportunities for optimization or broader application.

This study implemented the TRIZ-based reverse engineering approach to address the problem of accuracy decline in CNN models as the number of image classes increases. By systematically deconstructing existing solutions and analyzing contradictions between key engineering characteristics, we were able to define specific TRIZ principles to propose practical solutions.

The use of data augmentation, dimensional adjustment, class weighting, and transfer learning proved to be effective strategies for improving model accuracy. These methods, when viewed through the TRIZ framework, provided a structured approach to solving the technical contradictions between class quantity and accuracy. The reverse engineering perspective further enhanced this process by allowing us to identify hidden opportunities for improvement within the existing solutions.

While this study focused on CNN models, the principles of TRIZ could be applied to other machine learning architectures, such as recurrent neural networks, transformers, or generative adversarial networks. Future research could investigate how TRIZ can be adapted to solve contradictions in these more advanced architectures.

References

Alomar, K., Aysel, H.I., & Cai, X. (2023). Data augmentation in classification and segmentation: A survey and new strategies. *Journal Imaging*, 9, 46.

Atasever, S., Azginoğlu, N., Terzi, D.S., & Terzi. (2023). A comprehensive survey of deep learning research on medical image analysis with focus on transfer learning. *Clinical Imaging*, 94, 18–41.

Brad, S., & Brad, E. (2023). Using TRIZ to handle small datasets in artificial intelligence. In:

Acta Technica Napocensis Series Applied Mathematics, Mechanics, and Engineering, Technical University of Cluj-Napoca, Romania, p66.

Gadd, K. (2011). *TRIZ for Engineers: Enabling Inventive Problem Solving*. John Wiley and Sons, United States.

Guo, Y., Zhang, L., Hu, Y., He, X., & Gao, J. (2019). A survey on deep learning based face recognition. *Computer Vision and Image Understanding*, 189, 102805.

Kamnitsas, K., Ledig, C., Newcombe, V.F., Simpson, J.P., Kane, A.D., Menon, D.K., *et al*. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*, 36, 61–78.

Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., *et al*. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88.

Pan, S.J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.

Perez, L., & Wang, J. (2017). *The Effectiveness of Data Augmentation in Image Classification Using Deep Learning*; [arXiv preprint].

Sheu, D.D., & Lee, H.K. (2011). A proposed process for systematic innovation. *International Journal of Production Research*, 49(3), 847–868.

Shorten, C., & Khoshgoftaar, T.M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 60.

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A survey on deep transfer learning. In: *International Conference on Artificial Neural Networks*. Springer, Cham, p270–279.

Tribuana, D., & Arda, A.L. (2024). Image preprocessing approaches toward better learning performance with CNN. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 8 (1), 1–9.

Yang, Z., Sinnott, R.O., Bailey, J., & Ke, Q. (2023). A survey of automated data augmentation algorithms for deep learning-based image classification tasks. *Knowledge and Information Systems*, 65, 2805–2861.

AUTHOR BIOGRAPHY

Merve Cosgun completed her undergraduate degree in Industrial Engineering at Uludağ University. She is currently pursuing her Master's degree in Industrial Engineering at Bursa Technical University. Merve works as an industrial engineer at a private company and has a keen interest in the theory and applications of TRIZ.

Koray Altun is an assistant professor of Industrial Engineering at TU Bursa and a research collaborator at GLORAD. He is also the founder and manager of a startup focused on digital innovation and software consultancy. He holds a Ph.D and a BSc. Degree in Industrial Engineering from Gaziantep University and Erciyes University, respectively. He has published papers in well-respected journals and has led consulting engagements in reputable companies with a focus on R&D, technology, and innovation management. His recent research interests include the Innovation Excellence Model, Systematic Innovation, Technology and Innovation Management, and R&D Management.