

Attention Mechanisms-Based CNN Regression Model for Optimal Design of SPMSM

Tae-Hyuk Ji¹, Seok-Won Jung¹, *Member, IEEE*, In Seok Song¹ and Sang-Yong Jung¹, *Member, IEEE*

¹Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon 16419, Republic of Korea

This paper proposes an optimal design method for surface-mounted permanent magnet synchronous motors (SPMSM) based on electromagnetic performance prediction using a convolutional neural network (CNN)-based regression model enhanced with attention mechanisms. By capturing spatial structures and local patterns, CNNs effectively extract key features from motor designs, making them well-suited for performance prediction. The integration of attention mechanisms further improves the model by assigning greater weights to critical design features, enhancing prediction accuracy and learning efficiency. Additionally, attention mechanisms enable the identification of relationships among key elements, addressing the limitations of localized learning in conventional CNN models. In this study, the effectiveness of attention mechanisms is evaluated, and a CNN-based regression model incorporating these mechanisms is developed. The proposed regression model is then applied to optimize the design of an SPMSM.

Index Terms—Attention mechanisms, CNN, Permanent magnet machines

I. INTRODUCTION

Surface-mounted permanent magnet synchronous motors (SPMSM) are extensively utilized in high-performance drive systems owing to their remarkable torque density, high efficiency, superior dynamic response, and precise controllability. However, SPMSMs inherently suffer from torque ripple and cogging torque. This phenomenon not only compromises the output torque quality but also exacerbates noise and vibration, significantly impacting the overall performance of the machine. Consequently, the design of SPMSMs involves considerable complexity, as it must address multiple objectives and manage high-dimensional design variables.

To address these challenges, optimal design methodologies based on optimization algorithms have been extensively explored [1]. These methodologies often rely on finite element method (FEM)-based simulations, which, while accurate, are computationally expensive and time-consuming. To resolve these computational demands, recent efforts have focused on surrogate modeling techniques, to accelerate the optimization process. Additionally, artificial neural networks (ANNs) have emerged as powerful tools for approximate computation in parameter optimization tasks, achieving promising results in reducing the design cycle time [2]. More recently, convolutional neural networks (CNNs), a specialized deep learning approach, have been employed to construct regression models for the optimization of interior permanent magnet synchronous motors (IPMSMs) [3]. These studies have demonstrated that CNNs trained on cross-sectional images of stator and rotor geometries can effectively capture critical design patterns.

For PMSMs, accurately extracting the geometric features of the permanent magnets, which are among the most influential factors affecting electromagnetic performance, is essential. However, unlike IPMSMs, SPMSMs exhibit relatively minor geometric variations in permanent magnet shapes, making it

more challenging for CNNs to extract meaningful design features and achieve accurate regression modeling.

To overcome these limitations, this paper proposes a CNN-based regression modeling approach augmented with attention mechanisms for SPMSMs. The integration of attention mechanisms enables the model to assign greater weights to critical design features, enhancing both the learning efficiency and prediction accuracy. By leveraging these capabilities, we aim to significantly reduce the computational cost associated with SPMSM optimization while maintaining or improving design quality.

The proposed method is applied to optimize SPMSM designs using a surrogate-assisted optimization framework. The results demonstrate the effectiveness of the attention-based CNN approach in capturing subtle geometric features and improving the overall optimization process. This work provides a pathway for more efficient and accurate design methodologies for SPMSMs, contributing to the advancement of high-performance motor drives.

II. PROPOSED ATTENTION MECHANISM-BASED CNN MODEL

A. Convolutional neural network

The convolutional neural network (CNN) is a deep learning model that processes and learns from data through the convolution operation. This operation enables the model to identify specific patterns in the data, particularly by focusing on the relationships between nearby pixels. As a result, CNNs are capable of preserving spatial structures, such as positional relationships and patterns, throughout the learning process [4]. This ability makes CNNs highly effective at capturing spatial structures and local patterns in images, which is particularly beneficial for extracting features from motor designs.

However, CNNs primarily concentrate on learning relationships between adjacent pixels, which limits their capacity to capture interactions between distant pixels or complex structural relationships within an image. In the case of motor electromagnetic performance, this limitation becomes critical because the overall performance is determined by the

global interaction between the stator and rotor. Therefore, addressing this issue is essential. Additionally, because the sensitivity of a motor's geometric variables to electromagnetic performance varies, it is crucial to emphasize and effectively learn the most significant regions of the design.

B. Attention mechanism-based CNN

The attention mechanism is inspired by the way humans focus their attention, emphasizing or assigning weights to the most important parts of the given input. In neural networks, the attention mechanism enables models to concentrate more on critical information by assigning higher weights to specific parts of the input while reducing the emphasis on less relevant parts.

There are various types of attention mechanisms, and this paper specifically addresses channel attention, spatial attention, and self-attention mechanisms.

Channel attention calculates the importance of channels, which encode information such as shape, boundaries, and color within an image, and highlights the channels of higher significance. In the case of SPMSMs, channel importance is evaluated based on material characteristics that are distinguished by the specific shape or color of the permanent magnets or the core.

Spatial attention analyzes the spatial locations within a feature map and focuses on the most critical regions. For SPMSMs, the importance is determined based on the positions of the core and the geometric configuration of the magnets.

Self-attention assesses how all elements within the data relate to one another or are contextualized. By leveraging self-attention, the limitations of CNNs' localized learning can be effectively addressed.

Figure 1 provides an overview of the proposed attention mechanism-based CNN. The full paper will include a detailed explanation of the proposed architecture, along with an in-depth analysis of the functionality and effectiveness of each attention mechanism.

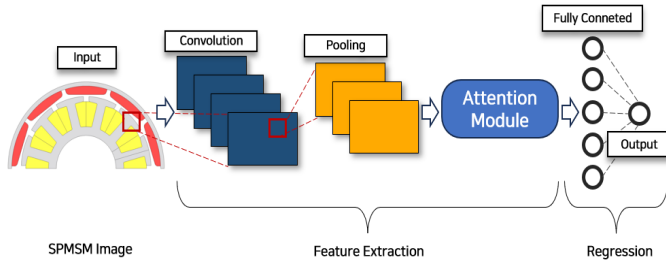


Fig. 1. Overview of the proposed Attention mechanism-based CNN.

C. Verification of the attention mechanism-based CNN

The CNN model used to verify the effectiveness of the Attention Mechanism, specifically ResNet 50, and the input data consisting of motor model shapes can be seen in Figures 1 and 2. A comparison was made between the model with both channel and spatial attention applied and the CNN model with no attention mechanism. The training conditions are provided in Table 1.

As shown in the comparison results in Table 1, the application of the Attention Mechanism led to a 40%

improvement in learning speed and a 32% reduction in the loss function. Additionally, no overfitting was observed during the training with the Attention Mechanism. The effects of each type of attention on the electromagnetic performance, aside from the average torque, will be discussed in detail in the full paper.

TABLE I
TRAINING CONDITIONS AND RESULTS

Item	Resnet 50	Resnet 50 + Attention
Prediction target	Average Torque	
Number of data	6,800	
Epoch	50	
Batch size	32	
Optimizer	Adam	
Learning rate	0.001	
Training trials	5	
Earllystop Epoch	27	16
Loss Funtion(MSE)	6.76e(-4)	4.57e(-4)

III. PROPOSED CNN MODEL-BASED OPTIMIZATION FRAMEWORK

The SPMSM is optimized using NSGA-II, a multi-objective optimization algorithm, with various electromagnetic performance metrics defined as objective functions. During the optimization process, the calculated data values are utilized to train the CNN regression model in the learning phase. Once the trained model achieves a sufficient level of accuracy, it is employed as a surrogate to partially replace the FEM in the optimization process. A comprehensive explanation of the detailed framework, the optimization algorithm, and the target performance will be provided in the full paper.

IV. CONCLUSION

This paper proposes an optimization design method for SPMSMs using a CNN model enhanced with attention mechanisms. The study investigates the characteristics and effects of various attention mechanisms in the context of SPMSM optimization and demonstrates their effectiveness. The application of both channel and spatial attention mechanisms resulted in a 40% improvement in learning speed and a 32% reduction in the loss function when optimizing for average torque. The full paper will further elaborate on the overall impact of the proposed approach in the context of optimal design using the optimization algorithm.

REFERENCES

- [1] X. Sun, N. Xu and M. Yao, "Sequential Subspace Optimization Design of a Dual Three-Phase Permanent Magnet Synchronous Hub Motor Based on NSGA III," in *IEEE Transactions on Transportation Electrification*, vol. 9, no. 1, pp. 622-630, March 2023
- [2] A. -C. Pop, Z. Cai and J. J. C. Gyselinck, "Machine-Learning Aided Multiobjective Optimization of Electric Machines Geometric-Feasibility and Enhanced Regression Models," in *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, vol. 4, no. 3, pp. 844-854, July 2023
- [3] H. Sasaki, Y. Hidaka and H. Igarashi, "Prediction of IPM Machine Torque Characteristics Using Deep Learning Based on Magnetic Field Distribution," in *IEEE Access*, vol. 10, pp.
- [4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436-444, May 2015