Fast Topology Optimization for PM Motors Using Variational Autoencoder and Neural Networks with Dropout

Hayaho Sato[1] and Hajime Igarashi[1]

[1]Graduate School of Information Science and Technology, Hokkaido University, Hokkaido 060-0814, Japan

This study proposes a novel topology optimization (TO) method for permanent magnet (PM) motors based on a variational autoencoder (VAE) and a neural network (NN). The VAE is trained to embed various shapes generated from the TO into the latent space. The NN is trained to predict the characteristics of the PM motor from its latent representation derived using the VAE. After training, TO is performed in the latent space based on the prediction using the NN. We adopt the Monte Carlo dropout to maintain prediction reliability using the NN during optimization, where prediction deviation is evaluated and used to eliminate unreliable solutions. The proposed method yields Pareto solutions within 80 s in a single-thread CPU machine, which is considerably faster than numerical analysis-based optimization, such as finite element analysis.

***Index Terms*—Design optimization, neural networks, permanent magnet (PM) motors.**

I. Introduction

SHAPE OPTIMIZATION, aimed at realizing high-performance electric machines in the industry, has attracted significant attention. Among these, the optimization of permanent magnet (PM) motors used in electric vehicles is becoming increasingly significant. Topology optimization (TO) has proven to be effective in the performance enhancement of PM motors, such as torque and efficiency. In TO, the magnetic core of the PM motor is modified by the generation and annihilation of holes. The PMs equipped in the rotor can also be modified using the multiphase TO method to represent more than three materials [1]. Moreover, a method exists in which the TO of the magnetic core and parameter optimization (PO) of PMs are hybridized [2]. However, performing such TOs requires hundreds to tens of thousands of numerical analyses using, for example, the finite element method (FEM), which is computationally expensive. Developing a TO method that can be performed with laptop computers in which every engineer in the industry can address is desirable.

To realize such a fast method, deep generative models, such as variational autoencoders (VAE) and generative adversarial networks, to design PM motors have been developed in recent years [3], [4]. Deep generative models are advantageous in that high-dimensional motor shapes (shape parameters or images) can be embedded into a low-dimensional latent space. After embedding the shapes, a surrogate model to predict their performance is built for the latent variables such that optimization can be performed in the latent space. This enables us to obtain optimal shapes without field computations at a low cost. Although these methods have been proven effective, the variety of motor shapes is restricted because they are generated based only on PO. Moreover, they can converge to the incorrect solution which are misjudged and having high performance by a surrogate model. This is because surrogate model-based optimization is often associated with regions of low prediction accuracy.

Manuscript received April 1, 2015; revised May 15, 2015 and June 1, 2015; accepted July 1, 2015. Date of publication July 10, 2015; date of current version July 31, 2015. Corresponding author: Hayaho Sato (e-mail: hayaho_sato@em.ist.hokudai.ac.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier (inserted by IEEE).

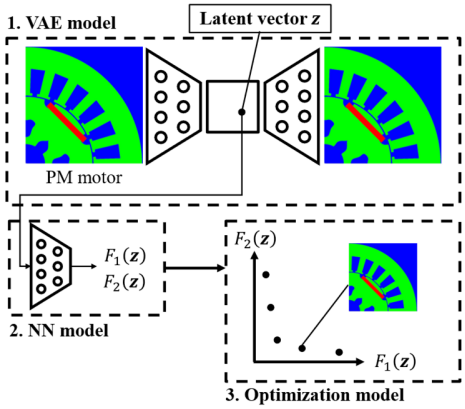


Fig. 1. Overview of proposed method.

This study proposes a novel TO method using a VAE and a neural network (NN) trained with dropout. The VAE model is used to embed various shapes generated from TO. The NN model is used as a surrogate model. To evaluate the reliability of NN-based predictions, we used Monte Carlo (MC) dropout [5], which estimates the standard deviation of predictions by sampling from NNs with different dropout states. The proposed method is applied to multi-objective TO of a virtual PM motor under speed-torque characteristic conditions.

II. Fast Topology Optimization Method

A. Overview of the Proposed Method

An overview of the proposed method is shown in Fig. 1. The proposed method comprises three models.

1. *VAE model.* This model embeds various images of motors generated from TO into a latent space via the encoder. Furthermore, it reconstructs the embedded images using the decoder. A latent vector, that is, an embedded motor image is represented as $\boldsymbol{z}$ hereinafter.
2. *NN model.* This model accepts $\boldsymbol{z}$ as input and predicts

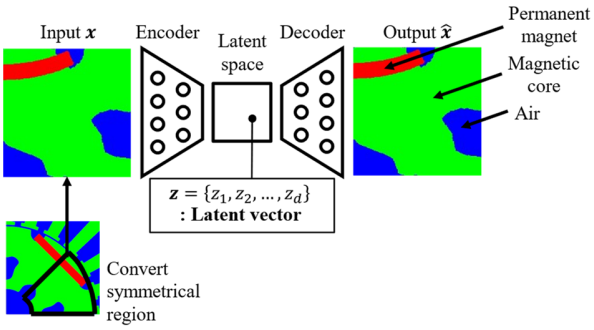


Fig. 2. VAE model to embed motor shapes into a latent space where d denotes the dimension of the latent space.

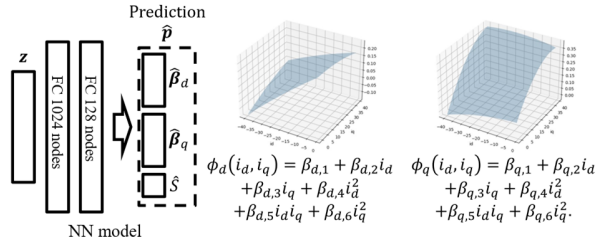


Fig. 3. NN-based model for the prediction of motor characteristics.

the characteristics of the corresponding motor shape. The prediction targets are the linkage fluxes of d- and q-axis, ϕ_d and ϕ_q, and the area of PM, S.

3. *Optimization model.* After training the VAE and NN models, we optimize the latent vector $\boldsymbol{z}$, which is associated with the motor shape, as explained. The objective functions $F_m(\boldsymbol{z})$ are evaluated only from the prediction of the NN model.

Because the dataset generated from TO is embedded and represented in the latent space, TO can be performed in the latent space using this method, which is cost-effective because field computations are not required and the dimension of the latent space is set smaller than that of the original space.

B. Variational Autoencoder for Embedding Shapes

The VAE model is shown in Fig. 2. The rotor of a PM motor is optimized in this study. The cross-sectional image of the motor is converted into a square image and used as the input $\boldsymbol{x}$. The VAE model accepts $\boldsymbol{x}$ and outputs the reconstructed image $\hat{\boldsymbol{x}}$ through the latent vector $\boldsymbol{z}$. The network architecture is based on that in [6].

In the VAE, prior $p(\boldsymbol{z})$ is assumed to have a standard normal distribution $\mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$. The encoder approximates the posterior to be a multivariate Gaussian with a diagonal covariance matrix, such that $\boldsymbol{z}$ is generated from the approximated posterior $q(\boldsymbol{z}|\boldsymbol{x})$:

$$\boldsymbol{z} \sim q(\boldsymbol{z}|\boldsymbol{x}) = \mathcal{N}\big(\boldsymbol{\mu}, \mathrm{diag}(\boldsymbol{\sigma})\big), \tag{1}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are the encoder outputs. Consequently, the loss function of the i-th input $\boldsymbol{x}^{(i)}$ can be derived as follows [7]:

$$L_{\mathrm{VAE}} = \left\|\boldsymbol{x}^{(i)} - \hat{\boldsymbol{x}}^{(i)}\right\|^2 + D_{\mathrm{KL}}\left(q\big(\boldsymbol{z}|\boldsymbol{x}^{(i)}\big)||\mathcal{N}(\boldsymbol{0}, \boldsymbol{I})\right), \tag{2}$$

where $D_{\mathrm{KL}}(\cdot)$ is the KL-divergence that regularizes $q\big(\boldsymbol{z}|\boldsymbol{x}^{(i)}\big)$ to be close to the prior, a standard normal distribution.

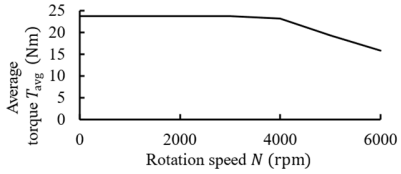


Fig. 4. Example of torque-speed characteristics of motors.

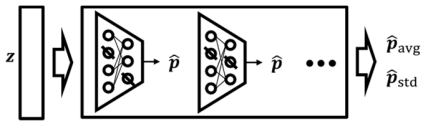


Fig. 5. Overview of MC dropout.

C. Neural Network for Performance Prediction

The NN model is shown in Fig. 3. The objective is to predict the d- and q-axis linkage fluxes of the PM motor, ϕ_d and ϕ_q, to calculate the average torque. The linkage fluxes depend on the d-and q-axis currents i_d and i_q. To represent this dependency, we approximate them using the quadratic forms [3], [8]:

$$\phi_d(i_d, i_q) = \beta_{d,1} + \beta_{d,2} i_d + \beta_{d,3} i_q + \beta_{d,4} i_d^2 + \beta_{d,5} i_d i_q + \beta_{d,6} i_q^2, \tag{3a}$$

$$\phi_q(i_d, i_q) = \beta_{q,1} + \beta_{q,2} i_d + \beta_{q,3} i_q + \beta_{q,4} i_d^2 + \beta_{q,5} i_d i_q + \beta_{q,6} i_q^2, \tag{3b}$$

whose coefficients are determined using the least-squares method. Consequently, the prediction targets of the NN model are the coefficients of the quadratic forms, $\boldsymbol{\beta}_d = \{\beta_{d,j}\}^T$ and $\boldsymbol{\beta}_q = \{\beta_{q,j}\}^T$ $(j = 1, \dots, 6)$. In addition, the area of the PMs, S, for optimization in the latent space should be evaluated. Finally, the output of the NN model is $\boldsymbol{p} = \{\boldsymbol{\beta}_d^T, \boldsymbol{\beta}_q^T, S\}^T \in \mathbb{R}^{13}$. The loss function of the NN model is

$$L_{\mathrm{NN}} = \left\|\boldsymbol{p}^{(i)} - \hat{\boldsymbol{p}}^{(i)}\right\|^2, \tag{4}$$

where $\hat{\boldsymbol{p}}^{(i)}$ denotes the NN model-based prediction.

D. Optimization Using Monte Carlo Dropout

After training the VAE and NN models, we optimize $\boldsymbol{z}$. The optimization problem is defined as follows:

$$\min. F_1 = -T_{\mathrm{avg}}(N_1), \tag{5a}$$

$$\min. F_2 = S, \tag{5b}$$

$$\mathrm{sub.\,to}\ T_{\mathrm{avg}}(N_i) - T_{\mathrm{avg}}^{\mathrm{target}}(N_i) \geq 0 \quad (i = 1,2), \tag{5c}$$

where $T_{\mathrm{avg}}(N)$ denotes the average torque at rotation speed N under maximum torque control, which is calculated as

$$T_{\mathrm{avg}}(N) = \max_{i_d, i_q}\left\{P_{\mathrm{n}}\big(\phi_d i_q - \phi_q i_d\big)\right\}. \tag{6}$$

where P_{n} is the number of pairs of poles which is 2 in this paper. An example of $T_{\mathrm{avg}}(N)$ is shown in Fig. 4. At a high rotation speed, the induction voltage $V \propto N$ reaches the limitation of the input inverter. To suppress this, flux weakening control is applied, which results in decreasing speed-torque

TABLE I
TORQUE CONDITION FOR PROBLEM I

Parameter	Value
Input current limitation (A)	40
Induction voltage limitation (V)	300
N_1, N_2 (rpm)	1000, 5000
$T_{\mathrm{avg}}^{\mathrm{target}}(N_1), T_{\mathrm{avg}}^{\mathrm{target}}(N_2)$ (Nm)	18.0, 15.0

TABLE II
SETTINGS FOR TRAINING

Parameter	Value
Dimension of the latent space	8
Optimization method	Adam
Initial learning rate*	0.001
Batch size	128
Number of epochs	100

*Exponential learning rate decay with a rate of 0.95 was applied.

characteristics. We impose constraints (5c) to maintain T_{avg} at a low speed N_1 and high speed N_2.

However, surrogate model-based optimization often misleads the individuals, which are misjudged as having high performance by a surrogate model. This is significant for VAE because misleading in the latent space may result in collapsed images. To prevent this and improve reliability of optimization, we adopted the MC dropout method to evaluate the reliability of the solutions [5]. An overview of MC dropout is shown in Fig. 5. In this method, NNs with different dropout states were randomly generated to make different predictions. This corresponds to sampling from the predictive distribution, where the weights of the NN model were regarded as stochastic variables. The average $\hat{\boldsymbol{p}}_{\mathrm{avg}}$ and standard deviation $\hat{\boldsymbol{p}}_{\mathrm{std}}$ were estimated from the sampled predictions. During optimization, the individual with an estimated $\hat{\boldsymbol{p}}_{\mathrm{std}}$ higher than t (%) of $\hat{\boldsymbol{p}}_{\mathrm{avg}}$ was eliminated by a heavy penalty:

$$F_m' = F_m + \alpha K \quad (m = 1,2), \tag{7a}$$

$$\alpha = \begin{cases} 1 & \text{if } \hat{p}_{j,\mathrm{std}} > t\left|\hat{p}_{j,\mathrm{avg}}\right| \text{ for any } j \\ 0 & \text{otherwise} \end{cases}. \tag{7b}$$

In this study, the number of samples, threshold t, and penalty K are set to 100, 30%, and 100. This strategy eliminates individuals for which the prediction of the NN model was highly variable and unreliable.

III. NUMERICAL RESULTS

A. Training Results of VAE and NN

The VAE and NN models were trained. The training data were obtained by preliminary solving the multi-objective optimization problem (5a), (5b), and (5c) using NSGA-II [9] under the torque conditions in Table I. During optimization, connection of the magnetic core was imposed by adding a penalty when it was separated. We obtained 32,052 training data from three optimization trials, where the type of PM was assumed to be I-, V-, and U-shaped. The image data resolution was set to 224×224 which is enough fine to represent the rotor shape. The VAE and NN models were simultaneously trained

(a) Input image (b) Reconstructed image

Fig. 6. Input and reconstructed images using the trained VAE.

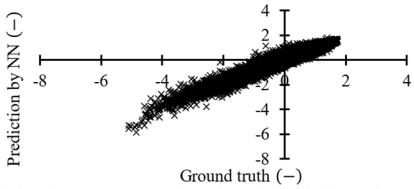


Fig. 7. Correlation between ground truth and prediction using the trained NN for p_1. Represented values are standardized.

[4], [10] using these data. The training settings are listed in Table II. The dimension of the latent space was determined to make the number of optimization variables as small as possible while keeping the accuracy in image reconstruction and regression by VAE.

The inputs and reconstructed images using the trained VAE are shown in Fig. 6. The inputs were successfully reconstructed. The correlation between the ground truth and prediction using the trained NN for $p_1 = \beta_{d,1}$ is shown in Fig. 7. The correlation coefficients for all targets $p_1, \ldots, p_{13}$ exceed 0.90, which would be sufficient for TO in the latent space.

B. Optimization Problem I

First, we consider the optimization problem, which is the same as the problem for training data collection, to validate the trained models. The parameters for NSGA-II are listed in Table III. In addition, the target torque in (5c) is set to 105% of $T_{\mathrm{avg}}^{\mathrm{target}}(N_i)$ to allow margin for the optimization [3].

The optimization results are shown in Fig. 8. For comparison, the training data were also plotted. The optimization in the latent space approximated the Pareto solutions of the training data, which included the U-shaped PM (i) and I-shaped PM (ii), (iii). Moreover, only 73.1 s was consumed in a single-thread machine, as shown in Table IV, which is considerably faster than FEM-based optimization.

To validate the Pareto solutions obtained in the latent space based on the NN model-based prediction, we converted the solutions to the mesh model, as shown in Fig. 9, and then analyzed them using FEM. The results are also shown in Fig. 8. The predicted and analyzed values had no significant gap. It is concluded that the NN model with MC dropout maintained the reliability of the solutions even when they were constructed from various shapes generated by TO. The analysis result for solution (ii) in Fig. 8 is shown in Fig. 10. This result satisfied the torque constraint (5c).

C. Optimization Problem II

The trained models can be applied to other optimization problems. Here, we considered the torque condition shown in Table V, which differed from the previous one. The input current limitation was reduced to 20 A. The target torque for each rotation speed was also changed.

The results are shown in Fig. 11. The Pareto solutions had U-shaped PM (iv), (v), and (vi). Moreover, we validated the Pareto

TABLE III
PARAMETERS FOR NSGA-II

Parameter	Value
Gene size	8
Population size	160
Crossover method	SBX
Number of children	80
Number of generations	50

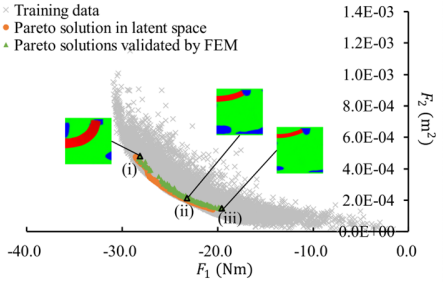


Fig. 8. Result of the multi-objective optimization (problem I).

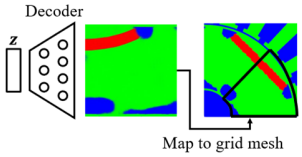
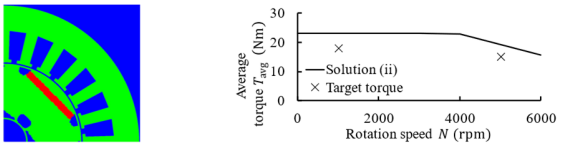


Fig. 9. Conversion of the latent vector to mesh, for FEM analysis.



(a) Shape of solution (ii) (b) Speed-torque characteristic

Fig. 10. Shape and speed-torque characteristics of solution (ii).

solutions using FEM. The Pareto solutions were successfully obtained using the proposed method even when the torque condition was changed from the original one. The accuracy of solutions with a small PM are quite low because the VAE model occasionally fails to reconstruct the image, as shown in Fig. 11, solution (vii). However, such anomalous solutions with extremely low torque can be easily excluded.

IV. CONCLUSION

This paper proposed a novel TO method based on VAE and NN. In this method, various motor shapes generated by TO can be treated. Optimization in the latent space yielded reliable solutions with a NN to which MC dropout was applied. We plan to extend the proposed method to consider double-layer PMs and synchronous reluctance motors. Moreover, we will develop a method to create the CAD data from the image obtained by the proposed method.

ACKNOWLEDGMENT

This work was supported by the JST SPRING (grant number JPMJSP2119), the MEXT Doctoral program for Data-Related InnoVation Expert Hokkaido University (D-DRIVE-HU) program, and KAKENHI 21H01301.

TABLE IV
COMPUTATION TIME

Stage	Computer specs	Computation Time
Data collection	CPU: 3.2 GHz, 64 threads CPU: 3.5 GHz, 16 threads	13.9 h
Training	GPU: NVIDIA Tesla V100 PCIE (16GB) x 2	4.0 h
Optimization in latent space (problem I)	CPU: 3.2 GHz, 1 thread	73.1 s
Optimization in latent space (problem II)	CPU: 3.2 GHz, 1 thread	55.6 s

TABLE V
TORQUE CONDITION FOR PROBLEM II

Parameter	Value
Input current limitation (A)	20
Induction voltage limitation (V)	300
N_1, N_2 (rpm)	1000, 5000
$T_{\mathrm{avg}}^{\mathrm{target}}(N_1), T_{\mathrm{avg}}^{\mathrm{target}}(N_2)$ (Nm)	9.0, 7.0

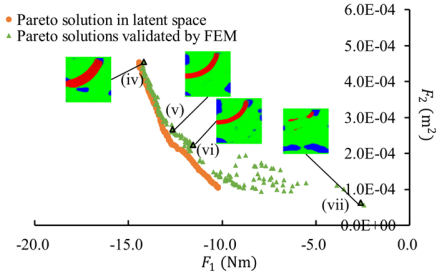


Fig. 11. Result of the multi-objective optimization (problem II).

REFERENCES

[1] T. Sato, K. Watanabe, and H. Igarashi, "Multimaterial Topology Optimization of Electric Machines Based on Normalized Gaussian Network," *IEEE Trans. Magn.*, vol. 51, no. 3, Art no. 7202604, 2015.

[2] S. Hiruma, M. Ohtani, S. Soma, Y. Kubota, and H. Igarashi, "Novel Hybridization of Parameter and Topology Optimizations: Application to Permanent Magnet Motor," *IEEE Trans. Magn.*, vol. 57, no. 7, Art no. 8204604, 2021.

[3] Y. Shimizu, S. Morimoto, M. Sanada, and Y. Inoue, "Automatic Design System with Generative Adversarial Network and Convolutional Neural Network for Optimization Design of Interior Permanent Magnet Synchronous Motor," *IEEE Trans. Energy Convers.* (early access), 2022.

[4] V. Parekh, D. Flore, and S. Schöps, "Variational Autoencoder-based Metamodeling for Multi-Objective Topology Optimization of Electrical Machines," *IEEE Trans. Magn.*, vol. 58, no. 9, Art no. 7500604, 2022.

[5] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning," *Proc. ICML 2016*, *arXiv:1506.02142v6 [stat.ML]*, 2016.

[6] B. Kim, V. C. Azevedo, N. Thuerey, T. Kim, M. Gross, and B. Solenthaler, "Deep Fluids: A Generative Network for Parameterized Fluid Simulations," *EUROGRAPHICS 2019, arXiv:1806.02071 [cs.LG]*, 2019.

[7] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," *ICLR 2014*, *arXiv:1312.6114v10 [stat.ML]*, 2014.

[8] T. Aoyagi, Y. Otomo, H. Igarashi, H. Sasaki, Y. Hidaka, and H. Arita, "Prediction of Current-Dependent Motor Torque Characteristics Using Deep Learning for Topology Optimization," *IEEE Trans. Magn.*, vol. 58, no. 9, Art no. 8205704, 2022.

[9] K. Deb, A. Pratap, and S. Agarwal, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 115–148, 2002.

[10] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, *et al.*, "Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules," *ACS Cent. Sci.*, vol. 4, no. 2, pp. 268–276, 2018.