those of the regular neural networks, which are trained with data from the target material only and initialized randomly using the default *PyTorch* tool.

The feed-forward neural network (FNN) is selected. FNN is one of the simplest and most commonly used neural networks [17], [18]. As shown in Fig. 5, the input layer of the FNN takes three inputs including frequency, flux density, and duty ratio of the excitation waveforms. The output layer of the neural network returns a single value, that is, the predicted core loss per unit volume. Logarithm values of the frequency, flux density, and core loss are used to train the neural network to better capture the quasi-exponential relationships [1]. The FNN has 3 hidden layers, whose dimensions are optimized using a hyperparameter-optimizing program *Optuna* [19]. Mean squared error is used as the loss function and *Adam* [20] is used for parameter optimization. The pretraining step itself is performed as regular neural network using the data from the first four materials. That is, there is no measurement data of the target material involved in the pretraining. A large amount of data for the four materials is available. The final value of each weight and bias after 75 epochs of pretraining (learning rate: 0.02) are recorded and this pretraining step is repeated for 10 times. Each parameter obtained are averaged over the 10 trials to achieve a stable starting point for the consecutive transfer learning. In order to show the effectiveness of the transfer learning, two networks of identical architecture are chosen, denoted as *Reference* and *Pretrained*. Both networks are initialized with two different sets of parameters - the Reference network is randomly initialized, and the Pretrained

Despite of the error, the result reveals the hypothesis that core losses of different materials, although different in values, share many patterns in common, and each has their unique features.

C. Re-training

The main motivation of transfer learning is to reduce the size of the dataset needed from the new material. The selected neural network is trained based on the available data for materials that are different from the target material, in order to capture the common characteristics. Afterwards a small dataset of the target material is used to re-train this pretrained neural network and include particular features. The final model can achieve comparable accuracy as if the neural network was trained by a much larger dataset from the target material.

Experiments are set up to examine the efficiency of the transfer learning process in achieving reasonable accuracy with a minimum dataset for training. Fig. 7 shows the results with a neural network that is trained with only 100 triangle data points from the target material without using pretrained models, and Fig. 8 shows the predicted core loss data by using the pretrained models after retraining them with the same amount of data. Fig. 9 shows the predicted core loss data by using a large amount of data from the target material without pretraining. All graphs were gathered with triangle 180 kHz excitations and three different duty ratios. The reference models that were trained by very few new data (Fig. 7) without pre-training are inaccurate. The pre-trained models that were given very few new data for retraining (Fig. 8) however, performed