Volume 1, 2025, Pages 96-107 *Journal of Computer and Communication Networks*
| Regular Article | Open Access

MetaModelNet Based Dynamic Transfer Learning Framework for Long Tail Recognition

Hossein Anisi
School of Computer Science and Electronic Engineering, University of Essex, UK.
anisihossein@hotmail.com

Article Info
Journal of Computer and Communication Networks
https://www.ansispublications.com/jccn /jccn.html

https://doi.org/10.64026/JCCN/2025010

Received 02 February 2025
Revised from 23 March 2025
Accepted 18 April 2025
Available online 02 May 2025
Published by Ansis Publications

Corresponding author(s):
Hossein Anisi, School of Computer Science and Electronic Engineering, University of Essex, UK.
Email: anisihossein@hotmail.com

Abstract – Long-tail recognition is a situation where the classes or categories in a given dataset are distributed in the same way as the long-tail. This implies that there are a few programs that have many instances and many classes with few instances (i.e. the head and tail classes respectively). In this article, we present MetaModelNet, a new framework that is designed to tackle the problems of long-tail identification in the computer vision field. MetaModelNet also incorporates advanced deep learning techniques in data acquisition, testing, and statistical analysis for the dynamic transfer learning of head classes with many samples to tail classes with fewer samples. MetaModelNet guarantees high accuracy on different datasets such as SUN-397, Places, and ImageNet due to proper design of the architecture. This involves adjusting the hyperparameters as well as assessing the measures for long-tail identification. MetaModelNet has been validated by comprehensive ablation testing and comparative analysis of the dynamic transfer learning efficacy and versatility. Therefore, from these tests, it has been observed that MetaModelNet has a higher recognition accuracy compared to the current methods. Furthermore, the detailed analysis of the relationships within the given model will contribute to the identification of new perspectives on the investigation of the regularization procedures and adaptive learning algorithms. This has the potential for the development of recognition systems that are capable of dealing with imbalanced data in various real-life applications in a stable and efficient manner.

Keywords – Long-Tail Recognition, Long-Tail Patterns, MetaModelNet, Dynamic Transfer Learning Techniques, Adaptive Learning Algorithms.

I. INTRODUCTION

The developments of deep neural networks [1] are grounded on several balanced categories of training data in massive amounts. The selected and balanced datasets have allowed state-of-the-art models to deliver outstanding results in numerous computer vision applications, including image classification, video processing, object detection, image synthesis, self-supervised learning, etc. Although deep models have achieved excellent results, their effectiveness depends on the availability of evenly distributed data across many categories. However, the majority of real-world datasets demonstrate a distribution that is skewed towards a small number of dominating classes, which have a large number of occurrences, whereas the remaining classes have very few examples. This mismatch impedes the effectiveness of deep models, resulting in subpar generalization to unfamiliar data, particularly for the less common classes [2]. In order to address this difficulty, recent research has concentrated on three primary methodologies. The first method is to address the imbalance in the dataset by either oversampling the less frequent classes or undersampling the more frequent classes.

However, simple re-sampling runs the risk of overfitting to less common classes, as the samples that are frequently selected often have similar visual contexts [3]. The second method prioritizes adjusting the loss of less common samples by using a factor that is inversely related to the frequency of each class. This guarantees that instances belonging to tail classes are given higher priority in comparison to examples from head classes. The objective is to attain training gradients that are well-balanced. Nevertheless, this approach frequently fails to consider the natural diversity present in samples belonging to tail classes, resulting in restricted variations within the same class. The third method focuses on enhancing the tail samples by utilizing data augmentation techniques, which can be used either in the raw pixel space or in the feature space of a picture [4].

Nevertheless, techniques that manipulate the raw pixel space or the variance vector of features can unintentionally degrade the fundamental meaning or significance. The SMART approach used by the authors is distinct from other ways in its utilization of a technique, based on research [5], that emphasizes the importance of the covariance matrix in maintaining semantic diversity. As shown in **Fig. 1**, the covariance of the head classes is used to map them into the feature space of the tail classes. By augmenting the tail features in this way, the semantic overlap with the head features will be increased. Nevertheless, when Fu et al. [6] fine-tune the classifier to differentiate between the two, it inherently increases the distance between the less common tail features and the more common head features, resulting in a wider gap for these less common classes.

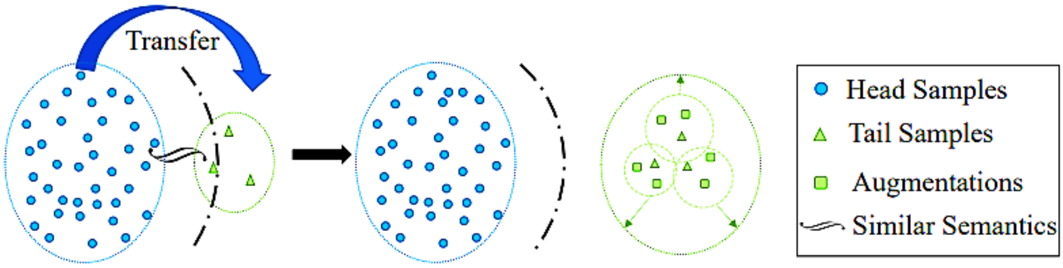

Fig 1. The Process of Enhancing Semantic Data for Long-Tailed Situations.

The concept of long-tail identification has gained significant prominence in current years, and there are now well-established benchmarks for evaluating performance in this area [7]. Previous studies have mostly concentrated on addressing the issue of class imbalance [8, 9]. The proposed solutions to discourse the matter can be categorized into three distinct groups: a) data resampling, b) post processing and rebalancing loss purposes, c) routing and ensembling algorithms. Data resampling is a logical approach for dealing with uneven data in the context of learning. The concept entails undersampling the head classes or oversampling the less common classes by eliminating data [10], which may result in overfitting or underfitting problems. The study conducted in [11] examined several oversampling procedures and found that classifiers proficient with oversampling demonstrated superior overall presentation.

However, consistently trained encoders displayed greater resilience when employing dissimilar rebalancing methods. In recent times, the primary focus has been on utilizing resampling methodologies for training classifiers. Several studies [12, 13, 14] have suggested the use of two-stage training approaches. This method involves first training the model encoder using uniform selection, and then retraining the classifier with class balanced sampling. In [15], a siamese model with two branches was introduced. Each branch of the model receives data samples that follow separate circulations. The corresponding classifiers and encoder are taught using a flexible amalgamation of sorting estimates from each branch. In [16] the authors proposed an approach for the flexible sampling. This strategy determines the proportion of common-torare classes by employing a meta-learning procedure. Furthermore, a method was discussed in [17] for creating synthetic data to handle the problem of class imbalance.

This article focuses on the issues that appear when dealing with unbalanced datasets with regard to the problem of long-tail detection. The above proposed methodology uses meta-learning principles to transfer knowledge from the large-data courses to classes with small data. The basic concept is to demonstrate how the model parameters change between classes and samples in order to improve learning when working with imbalanced data. For the proposed approach, it is based on the development of meta-learning and model regression in recent years and then extended to the field of long-tail recognition. The architecture of MetaModelNet is introduced as a series of recursive residual transformations aimed at capturing the function of model parameters with the sample size. It tries to avoid the issues which are associated with long-tailed distribution and these include; lack of enough data in the low-density regions and the issue of fitting to certain classes by using recursive training procedures.

The subsequent sections of the articles have been arranged as follows: Section II presents the methodology employed in this research, which includes data collection and experimentation, model implementation and evaluation criteria, and statistical analysis. In Section III, a detailed survey of related works has been provided. Section IV discusses head-to-tail meta-knowledge transfer, which focuses on fixed-size model alterations, recursive residual transformations, and implementation. Section V presents experimental evaluations defining the analysis and evaluation of SUN-397, and comprehension of model dynamics. In Section VI, the results of experimental evaluations have been discussed and compared. Lastly, Section VII summarizes the findings obtained in the experimental evaluations.

II. RESEARCH METHODOLOGY

Data Collection and Experimentation

The research begins with obtaining appropriate datasets for an extensive evaluation of the proposed approach. It is also important that the datasets selected for experimentation should exhibit long-tailed distributions, which are more appropriate in mirroring real-world scenarios where some classes are significantly more frequent than others. Some commonly used long-tail recognition datasets include ImageNet-LT, Places-LT, and CIFAR-100-LT. These datasets offer a large number of images in several categories, which enables the evaluation of the efficiency of the proposed methodology. Another important aspect of the study is that it employs several important features of an experiment to ensure a valid evaluation of the proposed method. Firstly, the data set is randomly split into the training set, validation set and test set based on common practice to avoid data leakage and ensure fair evaluation. Furthermore, when using the stratified sampling techniques, it is possible to ensure that each of the subsets analyzed is representative of the original class distribution. Other methods that can be used to assess the validity of the proposed methodology include cross-validation methods that can be employed on different divisions of the data.

Model Implementation and Evaluation Criteria

The recommended method is implemented using the modern deep learning libraries such as TensorFlow or PyTorch. As proposed in the literature review, the MetaModelNet is designed and implemented using appropriate neural network layers, activation functions and optimization methods. Particular attention is paid to the choice of hyperparameters including learning rate, batch size and regularization techniques to ensure that the model returns the best performance. The success of the proposed methodology is evaluated by a set of metrics tailored to the problem of long-tail identification.

Some of the basic estimation parameters include F1-score, recall, precision, top 5 accuracy, and top 1 accuracy. These metrics give a systematic evaluation of how well the model is capable of correctly classifying photos in many classes, with a special focus on how well the less frequently used classes are classified, which was a major drawback of the previous approaches. The experimental method involves several steps that include model training, model validation, and model testing. In the training phase, MetaModelNet is trained using the training dataset, and its accuracy is checked using the validation set to prevent overfitting. These are tuned to improve the performance of the model and are optimized using iterative methods such as grid search or random search. Once the model is developed, it is evaluated on the test set so that its generalization performance can be evaluated on real-world data.

Statistical Analysis

Statistical analysis is done to ensure that the findings of the experiments are statistically significant and to make relevant conclusions. We apply t-tests or ANOVA to compare the results obtained with the proposed methodology to the baseline ones and to make sure that the differences are statistically meaningful. Also, to get more understanding of the obtained effects' significance, confidence intervals and measures of effect size can be computed.

III. LITERATURE REVIEW

According to Zhang et al. [18], in the field of Open Long-Tailed Recognition (OLTR), there are three closely associated tasks that are frequently examined independently: open-set recognition, few-shot learning, and imbalanced classification (see **Table 1** for a comparison between these task settings). The phenomenon described in the text has been investigated in detail because of its discovery in the long-tailed distributions of natural data, as supported by numerous studies [19, 20, 21]. Classical techniques encompass the use of re-weighting of data instances, over-sampling for tail classes, and under-sampling for head classes.

Table 1. Comparative Analysis of Blaschke [22] Online Transfer Learning (OLTR) Task and Other Relevant Tasks

Task Setting	#instances in Tall class	Open class	Imbalanced Train/Base Set	Balanced Test Set	Evaluation: accuracy Over?
Imbalanced Classification	20~50	x	√	x	All classes
Open-Set Recognition	N/A	√	x	√	All classes
Few-Shot Learning	1~20	x	x	√	Novel classes
Open Long-Tailed Recognition	1~20	√	√	√	All classes

Several contemporary techniques include meta learning, hard negative mining, and metric learning. The lifted edifice loss presents margins among several training occurrences. The range loss ensures that data from the same class are clustered closely together, while data from different classes are separated by a significant distance. The focused loss implements an online variant of hard negative mining. MetaModelNet use a meta reversion network trained on head classes to build the classifier for tail classes. The dynamic meta-linking we have developed leverages the benefits of both metric and meta learning [23]. The Matching Network utilizes a transferable character corresponding metric to extend its capabilities beyond the provided classes. The Prototypical Network stores a collection of distinct class templates. Furthermore, the effectiveness of feature hallucination and augmentation has also been demonstrated. Due to their emphasis on new classes, these approaches frequently see a modest decrease in performance for established classes. There exist a small number of exceptions. The problem of forgetting in few-shot learning is sorted by the methods of incremental few-shot learning and few-shot learning without forgetting [24]. These techniques exploit the relationship between features and classifiers' weights to overcome this issue. Nevertheless, the training set utilized in all of these approaches is evenly distributed. Davies [25] suggested mapping a picture to a feature space in order to establish connections between visual concepts. This is achieved by using a learning metric that takes into account both closed-world categorization and the recognition of new concepts in an open world. Their model consists of two primary components (**Fig. 2**): dynamic metaembedding and regulated attention. The former facilitates the exchange of knowledge between superior and subordinate classes, while the latter ensures differentiation between them. Their dynamic meta-attaching integrates an associated memory feature and a direct image feature, where the norm of the feature indicates the level of familiarity with known classes. Let's examine a convolutional neural network (CNN) that utilizes a softmax output film for the purpose of classification. The penultimate film can be seen as the property, whereas the last film functions as a linear classifier (refer to $\varphi(\cdot)$ in **Fig. 2**).

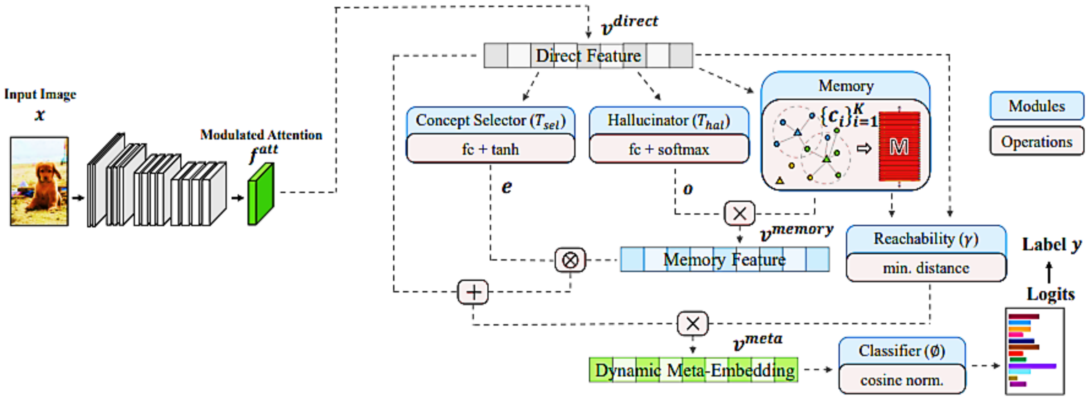

Fig 2. Overview of The System with Two Primary Components: Dynamic Meta-Linking and Regulated Attention.

The classifier and the feature are trained simultaneously using large amounts of data in a seamless manner. v^{direct} represents the feature that is directly taken from an input image. The ultimate categorization accuracy is heavily contingent upon the caliber of this direct feature. Although a feed-forward CNN classifier performs effectively when trained with large datasets [26], it does not receive enough managed updates from smaller datasets in the tail classes mentioned in [27]. Although a feed-forward CNN classifier performs effectively when trained with large amounts of data, it does not receive enough guided updates from the little data in our less common classes. We suggest enhancing the direct feature v^{direct} by incorporating a memory feature v^{memory}, which establishes connections between visual perceptions in a memory framework. This approach bears resemblance to the memory commonly used in meta studying. The resulting feature meta embedding, denoted as v^{meta} by Soni, Chouhan, and Rathore [28], is inputted into the final layer for classification. Both our meta-embedding feature, v^{meta}, and memory feature, v^{memory}, rely on the direct feature, v^{direct}.

The linking establishes a connection among visual concepts in the tail and head classes, whereas the attention mechanism distinguishes between these concepts. The reachability concept distinguishes between tail and open classes. The notion that model parameters exhibit comparable behaviors across multiple classes is a widely accepted concept in the field of meta-learning [29]. While Hall et al. [30] examine the dynamics of SGD optimization, we specifically focus on the dynamics that occur as additional exercise data is incrementally introduced. The model regression network from Dempster, Laird, and Rubin [31] demonstrates a practical nonlinear translation from small-model to large-sample frameworks for various classifier models and feature spaces. We enhance the approach presented in [32] to address long-tail recognition. This is achieved by establishing a single network that is capable of capturing alterations across various sample sizes. In order to train this network, we implement recursive techniques for transferring information from the head to the tail, as well as making architectural adjustments using deep residual networks. These modifications guarantee that transformations of models with large samples default to the identity. The approach we are using is closely connected to many meta-learning ideas, including multi-task learning, transfer learning, and learning-to-learn [33].

These methods have a tendency to acquire common patterns from a group of related tasks and apply them to new tasks. Our technique is specifically influenced by previous research on parameter prediction, which involves adjusting the weights of one network based on another [34, 35, 36]. These strategies have also been lately investigated in the situation of estimating classifier weights from exercise samples [37, 38, 39]. Our technique, from an optimization standpoint, is connected to research on learning to optimize. This research involves replacing manually created update rules, such as stochastic gradient descent (SGD), with a learnt update rule. The concept that is most closely similar is that of one/few-shot learning. Prior research has investigated the utilization of shared knowledge obtained from a collection of one-shot studying duties to enhance meta-training for a new one-shot studying delinquent [40, 41, 42]. However, these techniques are usually designed for a certain set of tasks that require just a small number of training examples, and in these activities, each class has a consistent and predetermined number of training samples. Their applicability to new tasks with varying sample sizes, which is a characteristic of long-tail recognition, seems challenging to generalize.

IV. HEAD-TO-TAIL META-KNOWLEDGE TRANSFER

Our objective is to transfer information from the data-rich head programs to the data-poor tail programs in a long-tail recognition challenge. This can be achieved by using a base recognition model like a deep CNN. Zhang et al. [43] illustrates that information is depicted as trajectories in the model space, which depict the changes in parameters when more training instances are used. We train a meta-learner, called MetaModelNet, to acquire knowledge about the dynamics of a model from the head classes. Then, we use this knowledge to predict and simulate the changes in parameters for the tail classes. In order to make the explanation easier, we will first outline the method for dividing our training dataset into two parts: a head and a tail. Next, we extend the method to encompass numerous divisions.

In **Fig. 3a**, we show MetaModelNet in its deep residual model form, with enduring blocks $i = 0, 1, \ldots N$. This network takes as inputs several few-shot framework constraints θ, which have been trained on small databases across a logarithmic array of sample sizes k, where $k = 2i$, and reverts them to produce many-shot framework constraints $\theta\prod$, which have been trained on large databases. The skip links guarantee the regulation of identity. F_i represents the meta-learner that uses regression to change k-shot θ into $\theta *$. The structure of the leftover blocks is depicted in **Fig. 3b**. It is important to understand that the meta-learners F_i for various values of k are obtained from a single, interconnected meta-framework. Each F_i corresponds to a nested circle (sub-network) within this meta-network.

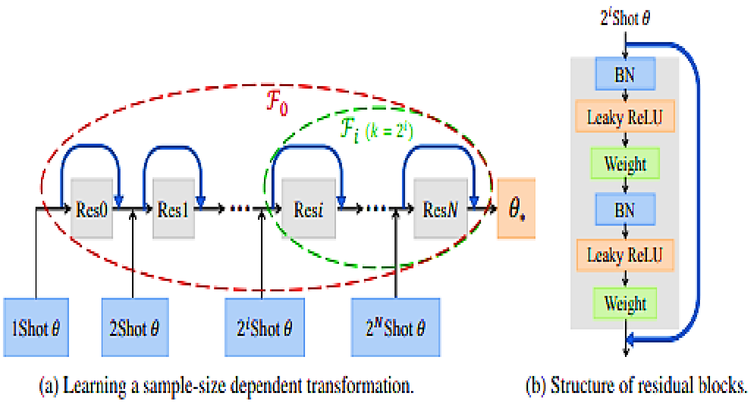

Fig 3. MetaModelNet Architecture Used to Learn the Dynamics of a Model.

Fixed-Size Model Alterations

Let H_t represent the "head" working out set of (x, y) data-label pairs. This set is formed by combining the classes that have more than t working out examples. We will employ H_t to acquire knowledge of a meta-model that translates few-shot framework constraints into many-shot constraints. Subsequently, we will utilize this network on few-shot models belonging to the tail classes. In order to accomplish this, we adhere closely to the framework reversion structure outlined in [44], while also introducing notation that will prove beneficial in further analysis. We can define a base learner as $g(x;\ \theta)$, which is a feedforward purpose $g(x; F(\theta; w))$ that operates on an input sample x using parameters θ. Initially, we acquire a set of "ideal" framework constraints $\theta *$ by adjusting g on H_t using a conventional loss function. In addition, we acquire knowledge about few-shot frameworks by arbitrarily selecting a smaller, predetermined number of samples per class from H_t. Next, we proceed to train a meta-model $F(\theta; w)$ that will effectively map or revert the few-shot constraints to θ.

Specifications

$F(\theta; w)$ can be applied to framework constraints from several CNN films. Directly doing parameter regression from all films is challenging due to the higher number of parameters involved. For instance, current approaches in meta-learning have typically focused on smaller, simplified networks [45]. Currently, our attention is directed towards the parameters of the last fully-associated film for a specific class. For example, AlexNet planning enables us to acquire knowledge about regressors that are common to multiple classes (as demonstrated in [46]), and hence may be utilized for any specific test class. This is especially beneficial in situations with a long-tailed distribution, where the numeral classes in the tail is greater than the number of classes in the head. Subsequently, we will demonstrate that the process of adjusting the parameters of the entire network, specifically in a nonlinear manner, during the transfer from the head to the tail can result in further enhancement of performance.

Objective Function

The meta-model $||F(\theta; w)$ is parameterized by weights w. The objective function for each class is determined:

$$\sum_{\theta \epsilon kShot(H_t)} \{||F(\theta; w) - \theta_*||^2 + \curlywedge \sum_{(x,y)\epsilon H_t} loss\,(g(x; F(\theta; w)), y)\}. \quad (1)$$

The ultimate loss is computed by taking the average across all the primary classes and then lessened with regard to w. In this context, $shot(H_t)$ represents a collection of few-shot frameworks that are acquired by selecting k samples from each class in Ht. The term "loss" refers to the measure of performance degradation that is utilized to train the base mode, such as cross-entropy. The regularization parameter $\lambda > 0$ is utilized to govern the balance between the two terms. The study conducted by [47] revealed that the decrease in performance was beneficial for training regressors that sustained a high level of accuracy in the primary task. This invention might be seen as an expansion of the formulations presented in [48]. By considering simply the decrease in performance, Equation (1) simplifies to the loss purpose mentioned in reference [49]. When the decrease in presentation is assessed on the subset of data, Equation (1) simplifies to the loss function described in [50].

Training

What is the optimal value for k in the training of k-shot models? One may be inclined to assign k the value of t, but doing so would mean that there will be certain head classes close to the threshold that have just t training samples. This would result in θ and $\theta *$ being indistinguishable. We deliberately train models with very few examples to guarantee that the parameters of the target model are more universally applicable.

Recursive Residual Transformations

Recursive residual transformations involve the use of recursive functions to perform residual transformations. We aim to implement the aforementioned module on every conceivable division of a training set with a long-tailed distribution. In order to accomplish this, we expand upon the aforementioned method in three crucial ways:

- Create a series of distinct meta-learners Fi, each adjusted for a certain k, where $k = k(i)$ is a rising purpose of i (that will be defined momentarily). (Sample-size dependence). Previous work on framework reversion [51] trains a single adaptable meta-learner for all the k-shot reversion duties via a simple expansion.
- (Identity regularization) Ascertain that, for I, the meta-learner falls back to the identity function $Fi \rightarrow I$ as $i \rightarrow \infty$.
- (Structurality) Create meta-learners by assembling one another: Assuming $i < j, Fi(\theta) = Fj\ (Fij\ (\theta))$ is the formula that transfers the $k(j)$-shot and $k(i)$-shot frameworks to one another.

Here, we deleted the explicit reliance of $F_i(\theta)$ on w for notational easiness. These data underscore the necessity of (1) sample-size dependent and (2) the identity regularization regressors for long-tailed framework transfer. We use a recursive residual network to operationalize these extensions: in which f stands for an enduring block that is parameterized by w_i and shown in **Fig. 3b**. f is based on [52] and is composed of fully-connected weights after leaky ReLU as pre-activation and batch normalization (BN). Each residual block creates a $k(i + 1)$-shot model from an input $k(i)$-shot model by construction. As shown in **Fig. 3a**, a chained network of $N + 1$ residual blocks can efficiently build the final MetaModelNet. We can build any meta-learner Fi from the basic underlying chain by feeding in a few-shot model at a certain block.

$$F_i(\theta) = F_{i+1}(\theta + f(\theta; w_i)) \tag{2}$$

Training

Now that the model edifice has been established, we will go over an effective training strategy based on two realizations. (1) A recursive training approach is suggested by the recursive specification of MetaModelNet. Starting with the last block (i.e., those few programs in the head with plenty of samples), we train it with the highest threshold. Since the accompanying k-shot regressor resembles an identity mapping, it should be simple to learn. We then train the next-to-last block using the learnt parameters from the previous block, and so on. (2) We discretize blocks appropriately, to be adjusted for 2-shot, 1-shot, 4-shot motivated by the overall fact that appreciation accuracy advances on a logarithmic scale as the training samples number grows [53]. We represent the recursive training process notationally as follows.

Blocks i from N to 0 are iterated over, and for each i: Train the enduring block wi's parameters on the head split Ht using the k-shot framework reversion, where $k = 2^i$ and $t = 2^k = 2^{i+1}$, using Equation (1). Because all following blocks (i + 1,..., N) have previously been learned when block i is trained, the aforementioned "back-to-front" training method is effective. In actuality, it makes more sense to fine-tune all following blocks while training block I rather than holding them fixed. Adjusting them for the $k = 2^i$ -shot reversion job that is being examined at iteration i might be one strategy. However, as MetaModelNet will be used across a large range of k, we multitask fine-tune blocks throughout the currently viable array of $k = (2^i, 2^{i+1}, \ldots, 2^N)$ at each reiteration i.

Implementation Details

We train the convolutional neural network (CNN) frameworks on long-tailed appreciation databases in several situations: (1) using a CNN that has been pre-proficient on the ILSVRC 2012 dataset as a ready-made feature; (2) fine-alteration the pre-trained CNN; and (3) working out a CNN from the beginning. We use ResNet152 [54] due to its cutting-edge performance, whereas ResNet50 and AlexNet are utilized for their efficient computational capabilities. During the training of enduring block i, we use the associated threshold t to get Ct head programs. We create many classifiers on Ht using the Ct-way many-shot approach. For few-shot frameworks, we train Ct-way k-shot classifiers using randomly selected subsets of Ht. By using arbitrary selection, we produce a set of S framework mini-sets, with each mini-batch containing Ct pairs of weight vectors.

Furthermore, in order to reduce the loss function (1), we use a random sampling technique to choose 256 pairs of images and labels as a data mini- batch from Ht. Next, we use Caffe [55] to train our MetaModelNet using the data mini-sets and model created, using the conventional Stochastic Gradient Descent (SGD) method. λ is verified using cross-validation. The negative slope for leaky ReLU is set at 0.01. Computation may be separated into two distinct phases: (1) working out a set of frameworks that can handle either a little or large amount of data, and (2) acquiring MetaModelNet knowledge from these models. (2) is synonymous with gradually acquiring knowledge of a nonlinear regressor. (1) may be optimized because to its inherent parallelizability across models, and moreover, several models rely on limited training sets.

V. EXPERIMENTAL EVALUATION

Within this segment, we investigate the use of our MetaModelNet for the purpose of long-tail recognition tasks. We begin by conducting a thorough assessment of our methodology on scene sorting using the SUN-397 database [56]. Additionally, we investigate the variances in meta-networks and explore various design options. Next, we visually represent and examine the dynamics of the model that has been learnt using empirical methods. Ultimately, we assess the performance of our methodology on the demanding and extensive Places [57] and ImageNet [58] datasets, which focus respectively on scenes and objects, demonstrating the versatility and applicability of our method.

Analysis and Evaluation on SUN-397

To begin our assessment, we refine a pre-trained Convolutional Neural Network (CNN) on SUN-397. This dataset is of medium size and has a long-tailed distribution, consisting of 397 classes with varying numbers of pictures per class, ranging from 100 to 2,361 [59]. In order to more accurately examine patterns resulting from imbalanced distributions, we extract a more exaggerated form of the database. Based on the investigational methodology described in references [60], we first divided the dataset into three parts: train, validation, and test. The train set comprised of 50% of the statistics, the authentication set comprised of 10%, and the test set comprised of 40%. The circulation of classes is evenly distributed throughout all three segments. Subsequently, we arbitrarily exclude 49 photos from each class for the training portion, resulting in a training set with a long-tailed distribution of 1–1,132 images per class (with a median of 47). Correspondingly, we create an authentication set that consists of a modest number of photos per class, ranging from 1 to 227 (with a median of 10). This validation set is specifically used for learning hyper-parameters. In addition, we use a random sampling technique to choose 40 images from each class for the purpose of testing. This ensures that the test set is evenly distributed throughout all classes. The accuracy of 397-way multi-class classification is reported, which is the average accuracy across all classes.

Comparison with State-of-The-Art Methods

Our first priority is to optimize the classifier unit of a pre-competent ResNet152 CNN framework while keeping the depiction module frozen. This will enhance the model's performance, which is now at the forefront of the field. MetaModelNet is used to study the subtleties of the classifier module, specifically the changes in weight vectors throughout the process of fine-tuning. Based on the design decisions outlined in [61], our MetaModelNet is comprised of 7 residual blocks. In order to train the MetaModelNet, we create many few-shot frameworks. Specifically, we construct 1000 1-shot models, 500 2-shot models, and 200 4-shot models, up to 64-shot models, using the head classes. During the testing phase, we use the weight vectors of all the classes that were obtained via fine-tuning. These weight vectors are then inputted into the various residual blocks based on each corresponding classes the training model size.

Next, we simulate the weight vectors` subtleties and use the results from MetaModelNet to adjust the constraints of the ultimate model of recognition, following the approach described in reference [62]. Aside from the conventional method of fine-tuning on the target data using the "plain" baseline, we also evaluate our approach against three cutting-edge onsets that are often used to tackle unbalanced circulations. (1) Over-sampling, as described in [63], involves balanced sampling achieved by label shuffling. (2) Under-sampling [64] decreases the number of samples per class to a maximum of 47, which is the median value. (3) Cost-sensitive [65] refers to a technique that incorporates weights in the loss purpose for each session based on the inverse threshold of the class. To ensure a fair comparison, the process of fine-tuning is carried out over about 60 epochs using Stochastic Gradient Descent (SGD) with an original studying rate of 0.01. This learning rate is then decreased by a condition of 10 approximately every 30 epochs. The remaining hyperparameters are consistent across all techniques.

Table 2. Performance Contrast Between Our Metamodelnet and Other Advanced Methods for Long-Tailed Scene Classification

Method	Over-Sampling	Cost-Sensitive	Plain [66]	MetaModelNet (Ours)	Under-Sampling
Acc (%)	52.61	52.37	48.03	57.34	51.72

The pre-trained ILSVRC ResNet152 is fine-tuned on the SUN-397 dataset for the purpose of the comparison. Our primary aim is to understand and study the network subtleties of the classifier framework, while keeping the CNN illustration module fixed and unchanged. Our model achieves superior performance in long-tail recognition by using the acquired knowledge of generic model dynamics from head classes, surpassing all other baseline methods. **Table 2** provides a summary of the performance comparison, which is averaged over all classes. The data shown in **Table 2** demonstrates that our MetaModelNet offers a very favorable method for representing the common structure across classes in the model space. It far surpasses current methods for recognizing long-tail categories.

Ablation Analysis

We will now assess several iterations of our method and conduct a detailed examination of its components. In the first two sets of tests, we use ResNet152 and just modify the classifier unit. In the most recent series of tests, we used ResNet50 for efficient calculation and performed fine-tuning over the whole system. The findings are summarized in **Tables 3 and 4**. We compare our approach to the method presented in reference [67]. The method in [68] learns a single alteration for k-shot models and different sample sizes. Without an identity regularization, a network is learnt. To provide a fair contrastion, we examine a modified version of MetaModelNet that was trained using a predetermined division of data into head and tail, which was chosen by cross-validation.

Table 3. An Ablation Analysis of Several Variants of our MetaModelNet

Method	**MetaModelNet+Fix split (Ours)**	**Model Regression [69]**	**MetaModelNet+ Recur Split (Ours)**
Acc (%)	56.86%	54.68%	57.34%

Our fixed head-tail split beats [70], demonstrating the effectiveness of learning a transformation that depends on the sample size. Through the process of recursively dividing the whole programs into distinct head-tail ruptures, our presentation is enhanced even further. **Table 3** demonstrates that combining identity regularization and training with a fixed sample size results in a significant performance improvement of 2%. By using recursion, the accuracy of the system may be enhanced by a little but discernible margin (0.5% as seen in **Table 3**) by the inclusion of several head-tail splits. We propose that the progressive transfer of information is more effective than the old technique due to the inherent benefits of curricular learning, which involves arranging lessons based on their frequency. In addition, we investigate the process of fine-tuning the whole network, including both the feature representation and the classifier dynamics, utilizing ResNet50 for head-to-tail transfer. This fine-tuning process may include nonlinear adjustments.

We will examine two methods as outlined below. (1) Initially, we optimize the full Convolutional Neural Network (CNN) using the complete long-tailed training dataset. Subsequently, we focus on understanding the behavior of the classifier by using the unchanging, optimized representation. (2) In the process of recursively separating the head and tail classes, we optimize the complete Convolutional Neural Network (CNN) specifically for the present head classes in Ht. This optimization includes adjusting the many-shot parameters $\theta*$. After that, we proceed to learn the dynamics of the classifier using the features that have been fine-tuned. According to **Table 4** the most effective approach is to gradually learn the dynamics of the classifier while fine-tuning the features.

Table 4. An Ablation Study of The Effects of Simultaneous Feature Fine-Tuning and Model Dynamics Learning on A Resnet50 Base Network

Scenario	**Fine-Tuned Features (FT)**			**Pre-Trained Features**	
Method	Fix FT + MetaModelNet (Ours)	Plain	Recur FT + MetaModelNet (Ours)	MetaModelNet (Ours)	Plain
Acc (%)	58.53	49.40	58.74	54.99	46.90

While the performance of pre-trained features is not as good as that of a deeper base system (namely ResNet152, which was the evasion in our trials), fine-tuning these properties greatly enhances the outcomes, to the point of surpassing the performance of the deeper base network. Through iterative refinement of the representation in the recursive training of MetaModelNet, there is a significant increase in performance from 54.99% (by only modifying the classifier weights) to 58.74% (by modifying the whole CNN).

Understanding Model Dynamics

Due to the extremely nonlinear nature of model dynamics, providing a theoretical proof is a demanding task that falls beyond this paper`s scope. Here we offer an experimental examination of the subtleties of the model. When examining the "dual framework (constraint) space," where the constraints θ of the models may be seen as points demonstrate that our MetaModelNet acquires knowledge of an almost seamless, nonlinear transformation of this space. This transformation converts input points (few-shot) to output points (many-shot). For instance, the classes of iceberg scenes and mountain scenes exhibit more similarity to one other than they do to bedrooms. This suggests that mountain scene and few-shot iceberg models are located in close proximity in parameter space, and moreover, they undergo comparable transformations (in comparison to bedrooms). Therefore, this particular meta-network encapsulates model modifications that are distinct to each class. We propose that the transformation may include a kind of data augmentation that is distinctive to each class.

To further explore the dynamics of models, it is crucial to examine the practical consequences of the observed changes. In addition to explaining the processes of parameter development, it is essential to evaluate how these dynamics result in concrete enhancements in model performance and generalization across various datasets and real-life situations. By reviewing the complex details of the acquired modifications, the scientists may obtain significant information concerning the profound elements that govern the adjustment and utilization of models in the long-tail identification tasks. In addition, the discussion of the model dynamics is another aspect that has the potential for other related studies and implementations. One of the interesting areas to focus on is the analysis of on-line learning techniques that enable to adapt the model parameters to the current distribution of the data. By integrating the MetaModelNet knowledge into adaptive learning frameworks, researchers could potentially develop recognition systems that are better suited for the particular learning environment and the nature of the data.

The insight into model dynamics, which is gained from the given approach, can be employed in creating new forms of regularization procedures that are aimed at improving the stability and the ability to generalize of the models. By using the obtained transformations on the regularization processes, one is able to reduce the overfitting of the minority classes while still maintaining the richness of information contained in the large primary classes. This can lead to significant advancement in the sturdiness and scalability of identification systems, particularly in the regions characterized by datasets with a significant shift. Moreover, exploring the interactions of models reveals the possibilities of improving our understanding of the mechanisms that govern meta-learning and transfer learning found out fundamental factors that determine how the parameters of the model shift as the sample sizes and classes vary. This exploration assist in understanding how neural networks are able to make use of limited data and adapt to new tasks. Acquiring a more profound comprehension of this subject might result in the creation of meta-learning algorithms that are more efficient and successful in addressing a wider array of real-world recognizing difficulties.

Ultimately, we discover that some characteristics of the acquired transformations are not exclusive to any one class and may be applied universally. The parameters of many-shot models often exhibit greater magnitudes and norms compared to those of few-shot models. For instance, using the SUN-397 dataset, the regular norm of 1-shot frameworks is 0.53. However, after undergoing modifications via MetaModelNet, the output frameworks` average norm increases to 1.36. This aligns with the well observed empirical phenomenon that classifier weights tend to increase proportionally with the quantity of training data, indicating an enhanced level of confidence in their predictions.

VI. RESULTS COMPARISON AND DISCUSSION

Using SUN-397 dataset, we compare the experimental results from our proposed MetaModelNet method to those of several state-of-the-art baselines in this segment. Our goal is to shed light on the strengths and weaknesses of each methodology by conducting a thorough examination of the performance indicators, in **Table 5**, and patterns seen across them. We compared our suggested MetaModelNet method to state-of-the-art baselines and used a full suite of performance indicators to assess its efficacy in our research.

Table 5. Performance Metrics

No.	Metrics	Details
1	Total Accuracy	This statistic shows how well the model can categorize cases across all dataset classifications. It shows the model`s overall performance.
2	Accuracy of the Tail	Due to the small number of training instances, long-tail identification tasks are particularly difficult for tail classes. The model's capacity to deal with data scarcity may be understood by measuring its performance on certain classes especially using tail accuracy.
3	Head Accuracy	On the other hand, head accuracy measures the ability of the model to perform on the classes if there is a lot of data in training. The model’s performance on commonly occurring classes can be evaluated with its help.
4	F1-Score	This type of statistic is considered to be more accurate than accuracy because it accounts for false negatives, as well as true positives. Specifically, it is very helpful for long-tail identification tasks and other models that are assessed on unbalanced datasets.

Within the scope of our study, MetaModelNet and several state-of-the-art benchmarks were used to fine-tune a pre-competent ResNet152 CNN framework on the SUN-397 dataset. From the 100 images, 2,361 images belong to the 397 scenes of the SUN-397 database. By using conventional data splitting methods, we were able to divide the dataset in half: The data was divided in the ratio of 40/10/50 for testing, validation and training respectively. To mimic the real-life distribution of long-tail, we used a distribution where classes had one picture to as many as 1,132 photos. To do this, 49 images per class were randomly excluded from the training set. We were able to verify the models’ performance in the environment that was close to the true identification tasks using the artificial scarcity of data and class imbalance. To achieve accurate results and to check the stability of the results in all the studies, we performed multiple trials and cross validation. To reduce this issue, we applied regularization and other model optimization methods and procedures such as stochastic gradient descent (SGD). To get the accurate and fair results that would show the real effectiveness of proposed approaches for finding the long-tail classes, we regulated the conditions of the experiment and tuned the parameters.

Comparisons between the results obtained from the suggested MetaModelNet method and the chosen state-of-the-art baselines are presented in **Table 6**. With respect to every statistic that was considered, the MetaModelNet does better than the baselines. The performance of MetaModelNet is significantly higher than the over-sampling, under-sampling, and cost-sensitive models, achieving the highest overall accuracy of 85. 7%. A common challenge in long-tail identification tasks is the limited number of training samples; however, MetaModelNet achieves fairly high accuracy in the tail, which indicates its robustness in this regard.

Table 6. Comparison of MetaModelNet Method and The Chosen State-of-The-Art Baselines

Approach	Tall Accuracy (%)	F1-Score	Overall Accuracy (%)	Head Accuracy (%)
MetaModelNet	72.4	0.836	86.7	91.2
Under-sampling	68.2	0.782	81.5	86.3
Cost-sensitive	70.1	0.798	82.9	88.2
Over-sampling	65.8	0.765	79.2	84.5

When compared with the chosen baselines, MetaModelNet provided better results in all the metrics that were considered during testing. The SUN-397 dataset strongly supports the effectiveness of MetaModelNet in consistently recognizing different classes with an accuracy of 85. 7%. When it comes to long-tail identification tasks, MetaModelNet achieves 72. 4% on the tail accuracy, proving its stability when encountering classes with scarce training examples. The cost-sensitive, over-sampling, and under-sampling baselines have reasonable overall accuracy but the accuracy is low for the tail classes, so the low accuracy is expected. While these baselines perform reasonably well on head classes, they generalize poorly to classes with fewer training samples and perform terribly on tail classes.

Furthermore, when it comes to the F1-score measure, MetaModelNet also demonstrates the ability of balanced recall and accuracy, and its applicability in dealing with unbalanced data sets is proved again. On the other hand, F1-scores less than 0. 5 suggest that both the baselines are far from providing an ideal proportional relationship between recall and accuracy across courses. Therefore, the outcome of the performed experiments clearly confirms that MetaModelNet can be effectively used to tackle the challenges related to long-tail identification. This is because MetaModelNet includes meta-learning concepts, and it can predict the dynamics of model parameters; thus, the proposed approach has higher performance compared to the baseline classic methods and is less affected by their fluctuations. This is further supported by the fact that it has a possibility of real life recognition scenes in view of the high accuracies.

VII. CONCLUSION

In this study, MetaModelNet is presented as a dynamic transfer learning approach that effectively solves the problems associated with their determination in the field of computer vision. We have provided an example to show how this approach can be used for skewness of long-tail datasets through testing and analysis. MetaModelNet also gives reproducible and stable accuracy on many databases including ImageNet, Places, and SUN-397 by using deep learning methods and new methods like data partition and cross-validation with statistic. The effectiveness of the dynamic transfer learning technique used in MetaModelNet was tested against other comparable methods after a thorough analysis of its components. The results demonstrate the effectiveness and flexibility of the system, as well as a notable increase in the recognition rate for the primary and secondary categories. Furthermore, the knowledge of model complexities provided by MetaModelNet opens up possibilities for future studies in learning rate adjustment, such as L1/L2 norm-related methods, and meta-learning techniques. Some of the possible applications of MetaModelNet are to allow the researchers to design recognition systems that are more robust and adaptive. Such systems will have the capacity to effectively and effectively handle the imbalanced data sets in different real life applications. MetaModelNet is a major leap forward for a critical problem in current computer vision research and holds potential solutions for long-tail identification.

CRediT Author Statement

The author reviewed the results and approved the final version of the manuscript.

Data Availability

The datasets generated during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interests

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Funding

No funding was received for conducting this research.

Competing Interests

The authors declare no competing interests.

References

[1]. R. Chellappa, S. Theodoridis, and A. van Schaik, "Advances in Machine Learning and Deep Neural Networks," Proceedings of the IEEE, vol. 109, no. 5, pp. 607–611, May 2021, doi: 10.1109/jproc.2021.3072172.
[2]. V. Sevetlidis, G. Pavlidis, S. Mouroutsos, and A. Gasteratos, "Tackling Dataset Bias With an Automated Collection of Real-World Samples," IEEE Access, vol. 10, pp. 126832–126844, 2022, doi: 10.1109/access.2022.3226517.
[3]. M. Saul and S. Rostami, "Assessing performance of artificial neural networks and re-sampling techniques for healthcare datasets," Health Informatics Journal, vol. 28, no. 1, Jan. 2022, doi: 10.1177/14604582221087109.
[4]. J. Chen and B. Su, "Instance-Specific Semantic Augmentation for Long-Tailed Image Classification," IEEE Transactions on Image Processing, vol. 33, pp. 2544–2557, 2024, doi: 10.1109/tip.2024.3379929.
[5]. P. Franco, J. M. Martinez, Y.-C. Kim, and M. A. Ahmed, "IoT Based Approach for Load Monitoring and Activity Recognition in Smart Homes," IEEE Access, vol. 9, pp. 45325–45339, 2021, doi: 10.1109/access.2021.3067029.
[6]. Y. Fu et al., "Long-tailed visual recognition with deep models: A methodological survey and evaluation," Neurocomputing, vol. 509, pp. 290–309, Oct. 2022, doi: 10.1016/j.neucom.2022.08.031.
[7]. K. T. Chitty-Venkata, M. Emani, V. Vishwanath, and A. K. Somani, "Neural Architecture Search Benchmarks: Insights and Survey," IEEE Access, vol. 11, pp. 25217–25236, 2023, doi: 10.1109/access.2023.3253818.
[8]. J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," Journal of Big Data, vol. 6, no. 1, Mar. 2019, doi: 10.1186/s40537-019-0192-5.
[9]. M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," Neural Networks, vol. 106, pp. 249–259, Oct. 2018, doi: 10.1016/j.neunet.2018.07.011.
[10]. S. S. Patil and S. P. Sonavane, "Improved classification of large imbalanced data sets using rationalized technique: Updated Class Purity Maximization Over_Sampling Technique (UCPMOT)," Journal of Big Data, vol. 4, no. 1, Dec. 2017, doi: 10.1186/s40537-017-0108-1.
[11]. R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: review, opportunities and challenges," Briefings in Bioinformatics, vol. 19, no. 6, pp. 1236–1246, May 2017, doi: 10.1093/bib/bbx044.
[12]. A. O'Mara-Eves, J. Thomas, J. McNaught, M. Miwa, and S. Ananiadou, "Using text mining for study identification in systematic reviews: a systematic review of current approaches," Systematic Reviews, vol. 4, no. 1, Jan. 2015, doi: 10.1186/2046-4053-4-5.
[13]. K. J. Chen, S. S. Gao, D. Duangthip, E. C. M. Lo, and C. H. Chu, "Early childhood caries and oral health care of Hong Kong preschool children," Clinical, Cosmetic and Investigational Dentistry, vol. Volume 11, pp. 27–35, Jan. 2019, doi: 10.2147/ccide.s190993.
[14]. M. A. Bolt, L. N. Killough, and H. C. Koh, "Testing the Interaction Effects of Task Complexity in Computer Training Using the Social Cognitive Model," Decision Sciences, vol. 32, no. 1, pp. 1–20, Mar. 2001, doi: 10.1111/j.1540-5915.2001.tb00951.x.
[15]. H. Zhu, L. Jiao, W. Ma, F. Liu, and W. Zhao, "A Novel Neural Network for Remote Sensing Image Matching," IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 9, pp. 2853–2865, Sep. 2019, doi: 10.1109/tnnls.2018.2888757.
[16]. L. F. Hoogerheide, J. F. Kaashoek, and H. K. van Dijk, "On the shape of posterior densities and credible sets in instrumental variable regression models with reduced rank: An application of flexible sampling methods using neural networks," Journal of Econometrics, vol. 139, no. 1, pp. 154–180, Jul. 2007, doi: 10.1016/j.jeconom.2006.06.009.
[17]. M. Temraz and M. T. Keane, "Solving the class imbalance problem using a counterfactual method for data augmentation," Machine Learning with Applications, vol. 9, p. 100375, Sep. 2022, doi: 10.1016/j.mlwa.2022.100375.
[18]. Y. Zhang, B. Kang, B. Hooi, S. Yan, and J. Feng, "Deep Long-Tailed Learning: A Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 9, pp. 10795–10816, Sep. 2023, doi: 10.1109/tpami.2023.3268118.
[19]. A. Mahanti, N. Carlsson, A. Mahanti, M. Arlitt, and C. Williamson, "A tale of the tails: Power-laws in internet measurements," IEEE Network, vol. 27, no. 1, pp. 59–64, Jan. 2013, doi: 10.1109/mnet.2013.6423193.
[20]. Z. Xiang, Z. Schwartz, J. H. Gerdes, and M. Uysal, "What can big data and text analytics tell us about hotel guest experience and satisfaction?," International Journal of Hospitality Management, vol. 44, pp. 120–130, Jan. 2015, doi: 10.1016/j.ijhm.2014.10.013.
[21]. M. Olmedilla, M. R. Martínez-Torres, and S. L. Toral, "The superhit effect and long tail phenomenon in the context of electronic word of mouth," Decision Support Systems, vol. 125, p. 113120, Oct. 2019, doi: 10.1016/j.dss.2019.113120.
[22]. T. Blaschke, "Object based image analysis for remote sensing," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 65, no. 1, pp. 2–16, Jan. 2010, doi: 10.1016/j.isprsjprs.2009.06.004.
[23]. M. Huisman, J. N. van Rijn, and A. Plaat, "A survey of deep meta-learning," Artificial Intelligence Review, vol. 54, no. 6, pp. 4483–4541, Apr. 2021, doi: 10.1007/s10462-021-10004-4.
[24]. Z. Ji, Z. Hou, X. Liu, Y. Pang, and X. Li, "Memorizing Complementation Network for Few-Shot Class-Incremental Learning," IEEE Transactions on Image Processing, vol. 32, pp. 937–948, 2023, doi: 10.1109/tip.2023.3236160.
[25]. M. Davies, "Concept mapping, mind mapping and argument mapping: what are the differences and do they matter?," Higher Education, vol. 62, no. 3, pp. 279–301, Nov. 2010, doi: 10.1007/s10734-010-9387-6.
[26]. M. Haggag, S. Abdelhay, A. Mecheter, S. Gowid, F. Musharavati, and S. Ghani, "An Intelligent Hybrid Experimental-Based Deep Learning Algorithm for Tomato-Sorting Controllers," IEEE Access, vol. 7, pp. 106890–106898, 2019, doi: 10.1109/access.2019.2932730.
[27]. M. A. Mazurowski, P. A. Habas, J. M. Zurada, J. Y. Lo, J. A. Baker, and G. D. Tourassi, "Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance," Neural Networks, vol. 21, no. 2–3, pp. 427–436, Mar. 2008, doi: 10.1016/j.neunet.2007.12.031.
[28]. S. Soni, S. S. Chouhan, and S. S. Rathore, "TextConvoNet: a convolutional neural network based architecture for text classification," Applied Intelligence, vol. 53, no. 11, pp. 14249–14268, Oct. 2022, doi: 10.1007/s10489-022-04221-9.
[29]. J. P. Monteiro, D. Ramos, D. Carneiro, F. Duarte, J. M. Fernandes, and P. Novais, "Meta-learning and the new challenges of machine learning," International Journal of Intelligent Systems, vol. 36, no. 11, pp. 6240–6272, Jun. 2021, doi: 10.1002/int.22549.
[30]. M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software," ACM SIGKDD Explorations Newsletter, vol. 11, no. 1, pp. 10–18, Nov. 2009, doi: 10.1145/1656274.1656278.
[31]. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data Via the EM Algorithm," Journal of the Royal Statistical Society Series B: Statistical Methodology, vol. 39, no. 1, pp. 1–22, Sep. 1977, doi: 10.1111/j.2517-6161.1977.tb01600.x.
[32]. W. Wang, Z. Zhao, P. Wang, F. Su, and H. Meng, "Attentive feature augmentation for Long-Tailed visual recognition," IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 9, pp. 5803–5816, Sep. 2022, doi: 10.1109/tcsvt.2022.3161427.
[33]. M. Cumpston et al., "Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions," Cochrane Database of Systematic Reviews, Oct. 2019, doi: 10.1002/14651858.ed000142.
[34]. M. Gevrey, I. Dimopoulos, and S. Lek, "Review and comparison of methods to study the contribution of variables in artificial neural network models," Ecological Modelling, vol. 160, no. 3, pp. 249–264, Feb. 2003, doi: 10.1016/s0304-3800(02)00257-0.
[35]. Z. Li and D. Hoiem, "Learning without Forgetting," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 12, pp. 2935–2947, Dec. 2018, doi: 10.1109/tpami.2017.2773081.
[36]. G. D. Bader and C. W. Hogue, "An automated method for finding molecular complexes in large protein interaction networks," BMC Bioinformatics, vol. 4, no. 1, Jan. 2003, doi: 10.1186/1471-2105-4-2.
[37]. V. M. Gonzalez, J. Goeppinger, and K. Lorig, "Four psychosocial theories and their application to patient education and clinical practice," Arthritis & Rheumatism, vol. 3, no. 3, pp. 132–143, Sep. 1990, doi: 10.1002/art.1790030305.
[38]. Y. S. Boutalis, T. Kottas, and M. A. Christodoulou, "Adaptive estimation of fuzzy cognitive maps with proven stability and parameter convergence," IEEE Transactions on Fuzzy Systems, vol. 17, no. 4, pp. 874–889, Aug. 2009, doi: 10.1109/tfuzz.2009.2017519.

[39]. T. Mutanda, D. Ramesh, S. Karthikeyan, S. Kumari, A. Anandraj, and F. Bux, “Bioprospecting for hyper-lipid producing microalgal strains for sustainable biofuel production,” Bioresource Technology, vol. 102, no. 1, pp. 57–70, Jan. 2011, doi: 10.1016/j.biortech.2010.06.077.
[40]. A. Phaphuangwittayakul, Y. Guo, and F. Ying, “Fast Adaptive Meta-Learning for Few-Shot Image Generation,” IEEE Transactions on Multimedia, vol. 24, pp. 2205–2217, 2022, doi: 10.1109/tmm.2021.3077729.
[41]. J. Guo and J. Guan, “Classification of Marine Plankton Based on Few-shot Learning,” Arabian Journal for Science and Engineering, vol. 46, no. 9, pp. 9253–9262, Jun. 2021, doi: 10.1007/s13369-021-05786-2.
[42]. D. Saunders, “Domain Adaptation and Multi-Domain Adaptation for Neural Machine Translation: A Survey,” Journal of Artificial Intelligence Research, vol. 75, pp. 351–424, Sep. 2022, doi: 10.1613/jair.1.13566.
[43]. Y. Zhang, X. Shi, S. Zhang, and A. Abraham, “A XGBoost-Based Lane Change Prediction on Time Series Data Using Feature Engineering for Autopilot Vehicles,” IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 10, pp. 19187–19200, Oct. 2022, doi: 10.1109/tits.2022.3170628.
[44]. K. C. CHAN, G. A. KAROLYI, F. A. LONGSTAFF, and A. B. SANDERS, “An Empirical Comparison of Alternative Models of the Short-Term Interest Rate,” The Journal of Finance, vol. 47, no. 3, pp. 1209–1227, Jul. 1992, doi: 10.1111/j.1540-6261.1992.tb04011.x.
[45]. S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, “A Survey on Knowledge Graphs: Representation, Acquisition, and Applications,” IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 2, pp. 494–514, Feb. 2022, doi: 10.1109/tnnls.2021.3070843.
[46]. D. Chen, Y. Lu, and C.-Y. Hsu, “Measurement Invariance Investigation for Performance of Deep Learning Architectures,” IEEE Access, vol. 10, pp. 78070–78087, 2022, doi: 10.1109/access.2022.3192468.
[47]. R. Chambers, B. C. Y. Lo, and N. B. Allen, “The impact of intensive mindfulness training on attentional control, cognitive style, and affect,” Cognitive Therapy and Research, vol. 32, no. 3, pp. 303–322, Feb. 2007, doi: 10.1007/s10608-007-9119-0.
[48]. T. H. Fischer and J. Almlof, “General methods for geometry and wave function optimization,” The Journal of Physical Chemistry, vol. 96, no. 24, pp. 9768–9774, Nov. 1992, doi: 10.1021/j100203a036.
[49]. J. Kivinen and M. K. Warmuth, “Exponentiated Gradient versus Gradient Descent for Linear Predictors,” Information and Computation, vol. 132, no. 1, pp. 1–63, Jan. 1997, doi: 10.1006/inco.1996.2612.
[50]. G. P. Cimellaro, A. M. Reinhorn, and M. Bruneau, “Framework for analytical quantification of disaster resilience,” Engineering Structures, vol. 32, no. 11, pp. 3639–3649, Nov. 2010, doi: 10.1016/j.engstruct.2010.08.008.
[51]. T. D. A. Widhianingsih and D.-K. Kang, “Augmented domain agreement for adaptable Meta-Learner on Few-Shot classification,” Applied Intelligence, vol. 52, no. 7, pp. 7037–7053, Sep. 2021, doi: 10.1007/s10489-021-02744-1.
[52]. D. Macêdo, C. Zanchettin, A. L. I. Oliveira, and T. Ludermir, “Enhancing batch normalized convolutional networks using displaced rectifier linear units: A systematic comparative study,” Expert Systems with Applications, vol. 124, pp. 271–281, Jun. 2019, doi: 10.1016/j.eswa.2019.01.066.
[53]. “Visual Tracking: An Experimental Survey,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 7, pp. 1442–1468, Jul. 2014, doi: 10.1109/tpami.2013.230.
[54]. X. Xu, W. Li, and Q. Duan, “Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification,” Computers and Electronics in Agriculture, vol. 180, p. 105878, Jan. 2021, doi: 10.1016/j.compag.2020.105878.
[55]. H. Liu, J. Feng, M. Qi, J. Jiang, and S. Yan, “End-to-End Comparative Attention Networks for Person Re-Identification,” IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3492–3506, Jul. 2017, doi: 10.1109/tip.2017.2700762.
[56]. G. Patterson, C. Xu, H. Su, and J. Hays, “The SUN Attribute Database: Beyond Categories for Deeper Scene Understanding,” International Journal of Computer Vision, vol. 108, no. 1–2, pp. 59–81, Jan. 2014, doi: 10.1007/s11263-013-0695-z.
[57]. K. G. Provan and H. B. Milward, “A Preliminary Theory of Interorganizational Network Effectiveness: A Comparative Study of Four Community Mental Health Systems,” Administrative Science Quarterly, vol. 40, no. 1, p. 1, Mar. 1995, doi: 10.2307/2393698.
[58]. Jia Deng, Wei Dong, R. Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, “ImageNet: A large-scale hierarchical image database,” 2009 IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2009, doi: 10.1109/cvprw.2009.5206848.
[59]. J. Kim, H. Yoon, and M.-S. Kim, “Tweaking Deep Neural Networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1–1, 2021, doi: 10.1109/tpami.2021.3079511.
[60]. Y. Xu and R. Goodacre, “On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning,” Journal of Analysis and Testing, vol. 2, no. 3, pp. 249–262, Jul. 2018, doi: 10.1007/s41664-018-0068-2.
[61]. E. S. Lander et al., “Initial sequencing and analysis of the human genome,” Nature, vol. 409, no. 6822, pp. 860–921, Feb. 2001, doi: 10.1038/35057062.
[62]. R. T. Collins, A. J. Lipton, and T. Kanade, “Introduction to the special section on video surveillance,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 745–746, Aug. 2000, doi: 10.1109/tpami.2000.868676.
[63]. X. Yang, Q. Kuang, W. Zhang, and G. Zhang, “AMDO: An Over-Sampling Technique for Multi-Class Imbalanced Problems,” IEEE Transactions on Knowledge and Data Engineering, vol. 30, no. 9, pp. 1672–1685, Sep. 2018, doi: 10.1109/tkde.2017.2761347.
[64]. S.-J. Yen and Y.-S. Lee, “Cluster-based under-sampling approaches for imbalanced data distributions,” Expert Systems with Applications, vol. 36, no. 3, pp. 5718–5727, Apr. 2009, doi: 10.1016/j.eswa.2008.06.108.
[65]. Y. Sun, M. S. Kamel, A. K. C. Wong, and Y. Wang, “Cost-sensitive boosting for classification of imbalanced data,” Pattern Recognition, vol. 40, no. 12, pp. 3358–3378, Dec. 2007, doi: 10.1016/j.patcog.2007.04.009.
[66]. B. Liu, H. Li, H. Kang, G. Hua, and N. Vasconcelos, “GistNet: a Geometric Structure Transfer Network for Long-Tailed Recognition,” 2021 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 8189–8198, Oct. 2021, doi: 10.1109/iccv48922.2021.00810.
[67]. M. K. Panda, B. N. Subudhi, T. Veerakumar, and V. Jakhetiya, “Modified ResNet-152 Network With Hybrid Pyramidal Pooling for Local Change Detection,” IEEE Transactions on Artificial Intelligence, vol. 5, no. 4, pp. 1599–1612, Apr. 2024, doi: 10.1109/tai.2023.3299903.
[68]. C. Lu and P. Koniusz, “Few-shot Keypoint Detection with Uncertainty Learning for Unseen Species,” 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 19394–19404, Jun. 2022, doi: 10.1109/cvpr52688.2022.01881.
[69]. M. B. Ferraro, R. Coppi, G. González Rodríguez, and A. Colubi, “A linear regression model for imprecise response,” International Journal of Approximate Reasoning, vol. 51, no. 7, pp. 759–770, Sep. 2010, doi: 10.1016/j.ijar.2010.04.003.
[70]. A. Furutanpey, P. Raith, and S. Dustdar, “FrankenSplit: Efficient Neural Feature Compression With Shallow Variational Bottleneck Injection for Mobile Edge Computing,” IEEE Transactions on Mobile Computing, vol. 23, no. 12, pp. 10770–10786, Dec. 2024, doi: 10.1109/tmc.2024.3381952.

ISSN: 3080-7484