https://doi.org/10.33472/AFJBS.6.11.2024.1314-1328

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

A HYBRID LEARNING MECHANISM TO ENHANCE THE FEW SHOT DATASET CLEANING

Vidya Mali[1], Dr. Shashidhar Gurav[2], Dr. Amit Chinchawade[3], Varsha Jujare[4]

[1]Department of Artificial Intelligence and Data Science, Sharad Institute of Technology College of Engineering Yadrav, Kolhapur, Maharashtra.
[2]Associate Professor, Department of Computer Science and engineering, Sharad Institute of Technology College of Engineering Yadrav, Kolhapur, Maharashtra.
[3]Associate Professor, Department Of Artificial Intelligence and Data Science, Sharad Institute of Technology College of Engineering Yadrav, Kolhapur, Maharashtra.
[4]Assistant Professor, Department of Computer Science and engineering, Sharad Institute of Technology College of Engineering Yadrav, Kolhapur, Maharashtra.

Email: [1]vidyamali2023@gmail.com, [2]gurav.shashidhar@gmail.com, [3]amitchinchawade@sitcoe.org.in, [4]varsha.jujare@gmail.com

Article Info

Volume 6, Issue 11, July 2024

Received: 23 May 2024

Accepted: 20 June 2024

Published: 09 July 2024

doi: 10.33472/AFJBS.6.11.2024.1314-1328

ABSTRACT:

In the supervised learning object or image identification process, it is necessary to train a large number of images to perform the prediction or search. It is a well-known fact that arranging such a large amount of data for supervised learning is not always feasible, and this insufficient data often leads to poor training results. When the data is insufficient, image classification employs various clustering and classification models like K-means and K nearest neighbors to overcome this challenge, which eventually works on the distance based techniques oftenly yield unsatisfactory results. Hence, few-shot learning models play a vital role in developing a well-trained model using a moderate or low amount of data. The augmentation process always powers few-shot learning, generating different views of an image and increasing the number of implicit data points to dynamically boost the image identification process. Hence, the proposed model utilizes the two convolution neural networks to train the datasets. Initially, we deploy a 2D convolution neural network on a large dataset to obtain the trained model. Subsequently, we train the few-shotdataset with a channel-boost convolution neural network to enhance the channel features of the selected images. The obtained two models tend to hybridize using channel factors to provide the best matching images for the testing input images. This research results are analyzed with some factors to unleash the difference between augmented and unaugmented training. Here, the proposed model provides good results in the initial criteria thereby improving the accuracy in Few-shot image cleaning.

Keywords: 2D Convolution Neural network, Channel boost Convolution neural network, few shot learning, Hybrid learning.

1. INTRODUCTION

Over the past few years, models that are based on deep learning have demonstrated remarkable success in tasks such as object detection and recognition of images. Several of these models are capable of performing at a level comparable to that of a person on difficult image classification datasets such as ImageNet, which has one thousand distinct object classifications. These models, on the other hand, are dependent on the supervised training paradigm, and the availability of labeled training data has a considerable impact on how well they perform. Additionally, the classes that the models are able to recognize are limited to the

ones that they were trained on.
As a result of the fact that there could not be sufficient enough tagged images for all classes throughout the training process, these models are less useful in situations that occur in the real world. Furthermore, we want our model to be able to detect images from classes that it was not exposed to during the training process. This is because it is nearly difficult to train on images of all possible objects just using images. Learning from a small number of examples is referred to as "few-shot learning," which is a difficulty.
One subfield that falls under the umbrella of machine learning is known as few-shot learning. In situations where there are only a few training samples with supervised data, it requires the categorization of new data. The performance of a computer vision model can be rather satisfactory even when just a limited number of training samples are used.
There are four distinct types of variations of Few-Shot Learning, which are generally recognized by experts. These variations include N-Shot Learning (NSL), Few-Shot Learning (FSL), One-Shot Learning (OSL), and Less than one or Zero-Shot Learning (ZSL).
The term "FSL" is typically used to refer to the N-way-K-Shot classification kind of classification. The letters N and K represent the number of classes and the number of samples per class that are used for training purposes, respectively. When compared to the other concepts, N-Shot Learning is considered to be the most comprehensive. Sub-fields of NSL are referred to as Few-Shot Learning, One-Shot Learning, and Zero-Shot Learning, respectively. Without any training examples, the goal of zero-shot learning is to classify classes that have not been seen before. We only have a single sample of each class when we use the One-Shot Learning method. When compared to OSL, Few-Shot is only a more adaptable variant because each class contains between two and five samples.
When it comes to tackling challenges with Few Shot Learning, there are generally two ways that we should take into consideration:

- Data-level approach (DLA)
- Parameter-level approach (PLA)

Data Level approach - The effectiveness of this method is not overly complicated. In order to construct a reliable model and avoid both underfitting and overfitting, it is based on the principle that additional data should be included in the event that there is insufficient data. As a result of this, a significant number of FSL problems can be solved by employing additional data from a substantial base dataset. One of the defining features of the original dataset is that it does not contain the classes that are included in our support set for the Few-Shot challenge. If, for example, we want to classify a specific species of bird, the base dataset might include photographs of a large number of other bird species.
Parameter level approach - A parameter-level perspective reveals that few-shot learning samples are relatively easy to overfit. This is due to the fact that these samples usually feature high-dimensional spaces that are vast in size. It will be possible to find a solution to this problem by restricting the parameter space, employing regularization, and making use of the proper loss functions. Despite the limited amount of training samples, the model will generalize effectively.
On the other hand, we are able to realize improvements in performance by directing the model to the extensive parameter space. There is a possibility that a conventional optimization strategy will not produce correct results since there is a paucity of training data.This is the reason why we train our model to find the optimal path across the parameter space in order to generate the most accurate predictions that are possible. This approach is commonly referred to as meta-learning.
[1] In this study, a new few-shot object Re-ID method called FSOR is introduced by Sheng-Hung Fan et.al. It efficiently builds object Re-ID models without requiring time-consuming annotation processes or arduous data gathering. Not only that, it ensures that object Re-ID

models can discriminate and generalize using an efficient few-shot learning model called H&C metric learning, which uses a re-parameterization method. The built-in object Re-ID model becomes more versatile and adaptable through the re-parameterization process, enabling it to re-identify objects that were not included in the training data. In order to improve the built model's discriminatory power, the suggested H&C metric learning merges the benefits of query-center distance and hard-mining distance. The experimental data presented by the author indicate that reparameterization and H&C metric learning can, on average, enhance mAP by over 17%. While building and running the FSOR model, the author also makes use of a number of straightforward but successful approaches, including data augmentation, a warmup learning rate, and label smoothing.
[2] In The TRX-IDP approach for few-shot action recognition was proposed by Yihang Ding et.al. and is based on the temporal Relational 530 CrossTransformers and the Image Difference Pyramid 531. TRX is the foundation of the author's 532 approach. This is why the question frame tuples 533 undergo sigmoid enhancement, scaling, and high-order image difference 534. It is possible to build the Image Difference Pyramid (IDP) 536 using motion feature information by combining it with the Motion 535 History Image (MHI). In addition to revising and improving the 539 model's linear mapping function, author 537 creates the CrossTransformers query representation for IDP 538. While TRX-IDP performs better than TRX on few-shot 540 benchmarks across all four datasets, it falls just short of HyRSM on Kinetics-400 and complete 543 SSv2, but attains state-of-the-art 541 performance on partial SSv2, HMDB51, and UCF101. The author plans to investigate other metric-based few-shot action detection algorithms 545 and attempt to merge them with the IDP module 544 in the future.
[3] An active sampling scenario for the few-shot classification algorithm has been provided by Junsuo Shin et.al, where the suggested system can readily improve performance. Taking into account the overall structure of several types of few-shot classification models, the system was created to be used in these models. A new way to build support sets that can represent the category's overall distribution was also introduced by the author: the distribution-to-distribution methodology. After conducting thorough experiments and analyses, the author identified the best combinations of system components for the goal few-shot classification method. The computational expense of the distribution fitting technique, however, is a limitation of the author's research. This issue will arise due to the handling of massive amounts of unlabeled data, however there are solutions that can help with it. In subsequent works, the author will address this matter.
Part 2 of this research study is devoted to reviewing the literature and conducting an in-depth analysis of prior efforts. Part 3 delves into the steps that are supposed to be taken. While Part 4 reveals the study's conclusions, In addition to wrapping up the research article, Section 5 provides chances for additional enhancements.

2. LITERATURE SURVEY

[4] Jing Bai et al. To learn a new class using a small amount of data and multiple iterations of fine-tuning, the authors of this article present the LPILC approach. The approach can quickly and efficiently adapt to a new class of HSI data without sacrificing computing expense. The primary responsibility lies with the incremental component's design. So that the incremental component can learn the new class and make use of the pretrained model at the same time, the author uses a linear programming model to adjust the pretrained model's classifier weight. Plus, the LPILC's architecture makes it possible for neural network models to pick up the new skill with just a single or small number of training examples, using very little time and

computational resources. Also, the LPILC can converge quickly and accurately after several epochs of fine-tuning. As far as the author is aware, this is the initial method to address incremental learning of subclass issues, offering a fresh viewpoint on resolving the new HSI classification challenge. The author applies the suggested LPILC to two frequent open-world setting challenges, 1) add-class and 2) subclass, to demonstrate the strength of her HSI classification approach. Hyperspectral remote sensing instances often have little training samples, so the author takes four possible outcomes into account to show how his suggested LPILC performs better and is more resilient when faced with data scarcity.

[5] A novel method for sequential or volumetric data segmentation was suggested by Dawood Al Chanti et.al., which combines the advantages of deep learning with expert interactions. Using a variety of techniques, including a Siamese network with subvolume recurrency, Bi-CLSTM, 3D ACS, and pseudo-labelling, the author takes use of the data's spatiotemporal coherence. Minimizing expert efforts during training, the resultant IFSSNet allows few-reference annotations to propagate over the full volume/sequence. The writer provided a comprehensive analysis of the US muscle segmentation and volume estimate challenge. This study aims, among other things, to validate the author's IFSS-Net using 3D freehand US volumes derived from youngsters suffering from Duchenne muscular dystrophy. Muscles are replaced by fatty tissues as the condition progresses, making segmentation more difficult. Therefore, a few adjustments seem to be in order. Training the network on a second domain that include adipose tissues or fine-tuning across a small population of patients with Duchenne muscular dystrophy could be two possible solutions. Alternatively, author may modify the popular zero-shot learning paradigm for classification and apply it to segmentation. Segmenting more anatomies that necessitate volume measurements and other sequential data-handling medical image processing activities could potentially benefit from the author's suggested methodology. Finally, there are two points of view: one is to see whether IFSS-Net can scale to handle numerous anatomies at once, and the other is to see whether it can generalize to different modalities.

[6] To improve few-shot object detection's data effectiveness, Jian Yao et.al. suggest DA-FSOD, a new data augmentation approach. In order to improve the data augmentation space, the author created a pool of operations based on numerous commonly used image processing procedures. In order to produce more varied enhanced versions while preserving the essential feature of the original input picture, the author then suggests a series and parallel connection system that layers the impacts of several procedures. Using the input image's semantic information, the author suggests imposed semantic data augmentation as a means to delve further into the deep feature information. The results of the experiments show that the author's method is superior to several common SOTA methods when it comes to ew-shot object recognition. In addition, the author's ablation investigations show that their method outperforms the conventional series and parallel connection settings, proving that the additional overhead of their proposed method is negligible in comparison to the performance benefit.

As the need for training larger and larger models grows, it becomes more challenging to clean datasets on this size, as described by M. B. Bijoy et.al. [7]. These massive models are vulnerable to the general performance and integrity threats posed by noisy data points. Dealing with outliers and noise in datasets is crucial as authors migrate to an ecosystem of systems driven by AI. Automated solutions to this problem are presented in this paper. A statistic driven cleaning strategy may be more suitable for situations with a more constrained dataset distribution, such as medical imaging. This method can also be applied to several domains of computer vision that face challenges with noisy data. It is demonstrated that representation learning approaches based on Deep Learning may automate the cleaning procedure using the results and visualizations provided above. Improving the model will need training it with a wider variety of datasets in future studies.

[8] Authors Viviane Clay et.al. have demonstrated that representations learnt by doing, without guidance from outside sources, store relevant information regarding the content of three-dimensional visual input. Semantic labels can be applied to these representations nearly as effectively as to representations that were optimized for object categorization using complete supervision. Like children's rapid mapping, this association works well with a small number of randomly selected labelled samples. Despite the lack of explicit instruction in semantic ideas, the author discovers that action-relevant objects, such as various types of doors, are encoded. Written in a style reminiscent of human learning, the author's results demonstrate the potential of a fresh method to train ANNs. This method first prioritizes unsupervised learning via interaction, and then moves on to supervised concept learning using sparsely labeled instances.

[9] As pointed out by Guangpeng Wang et.al., A new whole-classification model for few-shot picture identification using ViT is proposed by the author, who argues that all existing few-shot image classification approaches rely on meta-learning algorithms. This model can successfully extract features from the meta-train dataset and adapt to new class data. ViTFSL-baseline, a new, straightforward framework for few-shot image recognition learning, is proposed by the author by combining the whole-classification model with ViT and a unique NCL classifier. To facilitate the aggregation and classification of related features, the author uses ViT to train a large dataset and processes the features in the classifier. The author conducted thorough research on few-shot recognition tasks to prove that the ViTFSL-baseline they suggested gets good results. At the same time, the author shows how her approach, which uses simply the whole-classification method and the NCL classifier, is both successful and easy. Building on this work, the author plans to investigate the attention mechanism between the support set and the query set in future works. Additionally, it is expected that the model will exhibit strong performance in the semi-supervised few-shot learning situation while simultaneously being simplified.

[10] This work was conducted by Qingyang XU and colleagues. In order to increase the accuracy of Few-Shot human action identification, this study examines the Dynamic Time Warping (DTW) algorithm and its limitations. It then suggests ways to enhance its temporal alignment findings. In order to prove that the suggested algorithm improves the accuracy of Few-Shot human action detection tasks, experimental results were compared across various datasets and backbone networks.

[11] When it comes to diagnosing problems with wind turbines, in particular, the work of Farhan MD. Siraj et al. seeks to provide tangible benefits to the renewable energy industry by showing how state-of-the-art technology may contribute to environmental sustainability. New possibilities for resource-efficient error detection have emerged thanks to the author's suggested approach, which has allowed us to attain an impressive balance between diagnostic accuracy and computing efficiency. Furthermore, the experimental results demonstrate that the author's model reduces reliance on large labeled datasets, achieves competitive accuracy, and performs better than other options in terms of training time and resource intensiveness. Further highlighting its immediate significance and potential is its use for real-time fault detection on lightweight devices. There is an immediate demand for reliable fault diagnostic tools for wind turbines due to the expanding renewable energy industry. The author's work in this area is focused on improving the long-term viability and dependability of wind power generation.

Yan Huang et.al. [12] The authors of this article offer the Aligned Cross-Modal Memory (ACMM) as a solution to the understudied issue of fewshot image and sentence recognition. First, a weakly supervised cross-modal graph convolutional network for cross-modal regionword alignment; second, a persistently learned shared memory for cross-modal prototypical representation memorization; and third, adaptive fusion of multiple similarities using gated similarity fusion. the author has shown that their suggested model is effective by outperforming the state-of-the-art, and they have thoroughly examined how various components affect the final performance.

[13] A rapid detection method that can pinpoint the location of road landslides was introduced by Dat Tran-Anh et.al. The system was trained using a synthetic dataset that was built from the author's innovative design approach because real-world landslide photographs were difficult to acquire. In order to improve the system's ability to detect landslides in real-world scenarios, the author also suggested a novel few-shot segmentation method called CF-ASNet. The suggested CF-ASNet model and the produced dataset shown encouraging experimental results for road landslide damage classification and segmentation.

[14] According to Seo-Hyeong Park et al., there are many situations in which there aren't enough training samples, and time series data typically don't provide as much information as pictures. In order to deal with these few-shot circumstances, this paper suggests an FSL framework that trains the encoded pictures of the time series, TCA, and MFF. The author conducted experiments using the UCR archive to test the suggested strategy. The findings demonstrate that the method can produce a far better classification result while fusing rich representations and retaining information. The ablation study demonstrates that the author's framework is effective in its entirety, and that even the author's model performs well in situations where there are more than a few shots. The author's approach loses some information when scaling the encoded photos, which lowers its performance. As a result of data loss, this approach performs somewhat worse than the SOTA models. The author plans to broaden the few-shot scenario problem for multivariate TSC and further concentrate on improving performance without this information loss in future work. In order to expand the method's applicability in many areas, the author will also apply it to different kinds of time series, like unevenly time sampled time series.

[15] By Seunghun Lee et al. For better performance of text-to-image generative models and better intent capture, this study suggests a quick optimization approach. In the studies that included different tasks and phrase lengths, four-shot in-context learning performed better, especially when the text prompts were short. In terms of removing noise and including crucial keywords, the suggested prompt optimization methodology beats others, leading to more accurate and aesthetically pleasing image output, when compared to traditional methods like basic phrase summarizing. The results are important because they may have implications for many other areas where rapid optimization using big PLMs could be useful. The suggested method opens up new avenues for intelligent image generating systems by letting users customize the process and get the desired visual outcome.

Proposed Model

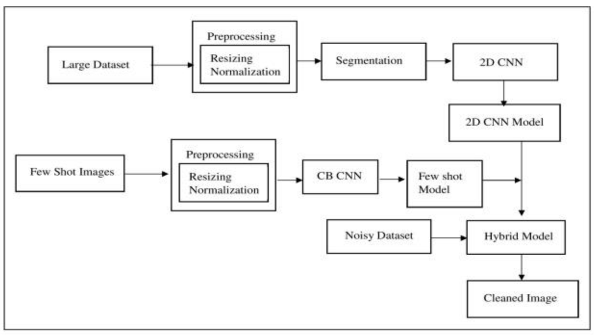

Figure 1: Overview of the proposed model for Few Shot image cleaning

The above figure 1 indicates the idea that is being used while building the few-shot image cleansing process. The steps are indeed narrated in depth in the below-mentioned phases.

Phase 1: Dataset Collection and Preprocessing – At this stage of the model's development, we get a large dataset from the following URL: https://www.kaggle.com/datasets/emmanuelpintelas/few-shot-datasets. A wide variety of image types, including birds, cars, characters, and more, are available at this URL. This URL has eleven distinct categories of color images that can be used for testing and training purposes as mentioned in [16],[17].

They begin the process of resizing and normalizing the images once they have been obtained. The first step in building the model is to sort the images into two categories: training and testing. Then, the proposed model is fed with these folders to train. All of the images in the collection are scaled to 128 × 128 pixels. Prior to training with a numpy array, all of the resized images are identified and mapped. In this way, images are given binary labels that allow for fast and accurate processing.

Phase 2: Data Segmentation – During this deployment phase, the ImageDataGenerator class constructs images with a rotation range of 40, a width shift range of 0.2, a height shift range of 0.2, a recall factor of 1/255, a shear range factor of 20%, a zoom range of 20%, a horizontal flip state of TRUE, and a fill mode of nearest. Following this process, we convert all the images into a numpy array in Python, reshaping them to fit the given height X width and color channel number of 3, as we are working with the RGB color image channel. After that, we'll segment all images in the training and testing sets using binary float 32 mode. This checks that the data in the numpy array, which is a binary array containing both the test and train images, is appropriately assigned to the selected category. The X[] and Y[] lists utilized for training and testing are generated by transforming labels into binary numpy arrays, in addition to the image data. As previously stated in the upcoming implementation phase, we are now ready to train the model using the 2D Convolution neural network mode.

Phase 3: Construction of 2 Dimensional Convolution neural network Model- The segmented images are now prepared to train the 2D convolutional neural network (CNN) model in this stage of the proposed model. The first step in this phase is to build an object of the neural network with the type sequential. This is because the neural network is supposed to traverse the layers in a linear fashion. Following this, the neural network's first layer is configured with 32 kernels, each having a 3 x 3 size, an input shape of 128 X 128 pixels, and 3 color channels. The activation function 'Relu' is applied after this initial layer; it consistently returns the maximum value of the data. The neurons are loaded into a 2 X 2 Max pooling layer after the 'Relu' activation function is applied.

The second layer also includes 32 kernels measuring 3 x 3 without an input shape, an activation function called 'Relu', and a Max pooling layer with a pool size of 2 x 2. Then, finally, the third contains 64 kernels, each with a dimension of 3 x 3 with the activation function ‘Relu’ and a max pooling layer with a pool size of 2 x 2. Following these 3 layers of convolution layers, the model is ‘Flattened’ with a dense layer of size 64 and with an activation function of ‘Relu’ again. A drop-out percentage of 50 is applied to the dense layer dimension 1 to end this with the activation of 'Sigmoid. ’

The optimizer 'rmsprop' evaluates the designed architecture of 2D CNN for the mode 'binary cross entropy', and then sets a batch size of 32 to train the model for 50 epochs, storing the trained data in a file with an extension. h5. The ‘Relu’ and'sigmoid’ activation functions can be seen in the below-mentioned equations 1 and 2 respectively.

$f(x) = \max(0, x)$ ____ (1)

Where, x is any positive value

$S(X) = \frac{1}{(1+e^{-x})}$ _______(2)

Where,
X is the input to a neuron
$f(x)$ = Relu Activation Function
$S(x)$ = Sigmoid Activation Function
e= Euler's Number

The Architecture of 2D CNN can be seen in the below mentioned figure 2.

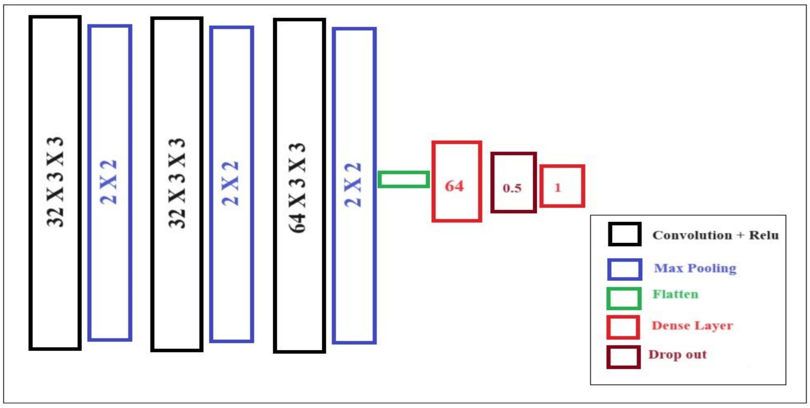

Figure 2: Architecture of 2D CNN Model

Phase 4: Constructing Channel Boost Convolution Neural network model – This step of the proposed approach involves building an image data generator object with a 1/255 aspect ratio and then conducting extensive analysis with the software libraries Keras and TensorFlow. Implementing training method on few shot images is the goal of the current methodology. There are a number of parameters that must be specified when creating a new instance of the ImageDataGenerator object. These include the locations of the training and testing folders, the image dimensions, a batch size of 32, and the categorical class setting, with grayscale being the specified color mode.In algorithm 1, we subject each pixel to an estimate of its RGB average, then set this estimated average value to the same pixel to obtain the grayscale features.

ALGORITHM 1: Absolute Gray scaling
// Input: Resized Image R_{IMG}
//Output: $GRAY_{IMG}$
// function: absoluteGrayScaler(R_{IMG})
1: **Start**
2: $ROI_{IMG} = \emptyset$
3:***for*** i = 0 to size of Width of R_{IMG}
4:***for*** j=0 to size of Height of R_{IMG}
5: color[] = $R_{IMG\,[i,\,j]}$ RGB
6: R= color[0]
7: G= color[1]
8: B= color[2]
9: ***AVG=(R+G+B)/3***
10. $GRAY_{IMG[i,\,j]}$ =[***AVG, AVG, AVG]***
11: ***end for***
12: ***end for***

13: ***return*** $GRAY_{IMG}$
14: ***stop***

The Sequential class in the TensorFlow kit is used to provide a framework for a sequential neural network. Afterwards, the "ReLU" activation function and a convolution layer with 32 3x3 kernels are integrated into the topmost layer of the deep neural network, which is devoted exclusively to the image equivalent measurements. This is followed by the incorporation of a second Convolutional layer with 64 3 x 3 kernels and the activation function of the rectified linear unit (ReLU). The implementation includes a pooling layer with a 25% dropout rate and a maximum size of 2 x 2.
Incorporating a third convolutional layer, the model makes use of 128 3 X 3 kernels with the ReLU activation function . The Max pooling layer is required to be 2 x 2. The fourth layer is added after the third layer has been deployed. It uses the Rectified Linear Unit (ReLU) activation function and has 128 3 x 3 kernels. The model design includes a subsequent Max pooling layer, a 2 by 2 dimensions, and a 25% dropout rate.
Incorporating a dense layer of 1024 units, triggered by the rectified linear unit (ReLU) function, and applying the flatten function constitute the final structure of the neural network. A 50% dropout rate is defined at the conclusion of the convolutional neural network using two dense layers and the "softmax" activation function. Softmax activation function is shown in the equation 3.

$$\sigma(z)i = \frac{e^{zi}}{\sum_{j=1}^{k} \quad e^{zj})} \text{\_\_\_\_\_\_\_}(3)$$

Where
$\sigma = \text{Softmax}$
$z = \text{Input Vector}$
$K = \text{No. of classes}$
$e^{zi} = \text{Standard exponenetial funtion for input vector}$
$e^{zj} = \text{Standard exponenetial funtion for output vector}$

During a 50-epoch training period, the Adam optimization technique is commonly used to improve the output. The trained data is saved in a format called.H5 once the training procedure is finished. The architecture of the Channel boost Convolutional Neural Network is shown in Figure 3.

Layers	Activation
CONV 32 x 3 x3	Relu
CONV 64 x 3 x3	Relu
MaxPooling 2 X 2	
DropOut 25%	
CONV 128 x 3 x3	Relu
MaxPooling 2 X 2	
CONV 128 x 3 x3	Relu
MaxPooling 2 X 2	
DropOut 25%	
Flatten	
Dense 1024	
DropOut 25%	
Dense 2	Softmax
Adam Optimizer	

Figure 3: Architecture of CB- CNN Model

Phase 5: Hybrid model to get cleansed data – This is the final phase of the system that has been proposed; in this phase, the model accepts an image from the user and then tests it against the trained data of the 2D CNN and the CB-CNN model. The trained data is used to obtain all of the images that are closest to the fed image. In order to obtain the few shots for the fed image, the characteristics of the huge dataset are initially retrieved by the.h5 procedure of the 2D CNN. After the features have been acquired, they are converted into numpy binary matrices so that they can be hybridized with CB CNN model in order to increase the channel values . As a result of this the process of compiling the few shot images and displaying them to the user in an interactive user interface is eventually accelerated.

3. RESULTS AND DISCUSSIONS

Python is used as the programming language for the machine that is used to deploy the proposed model that was produced for cleaning the few shot images. The machine has 8 GB of RAM and a Core i5 processor. In the process of identifying images using a few shots, the model is evaluated to determine the accuracy that it has attained. Both figure 4 and figure 5 illustrate the training accuracy for the few shot images that were captured using the 2D CNN and CB CNN models, respectively.

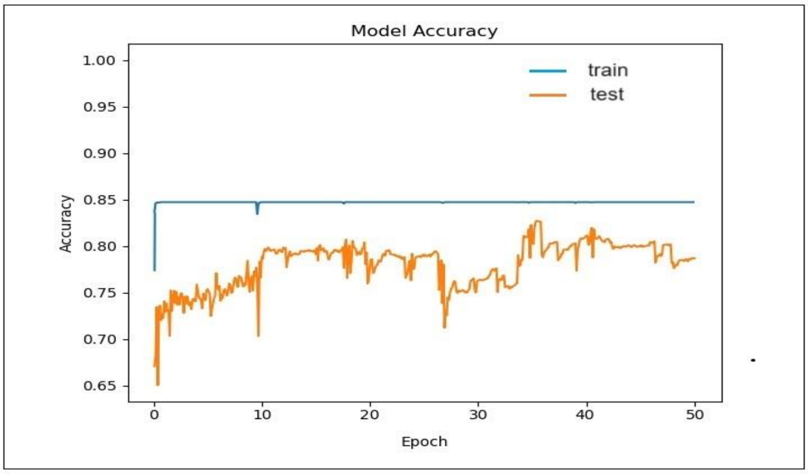

Figure 4: Trained accuracy for 2D- CNN model

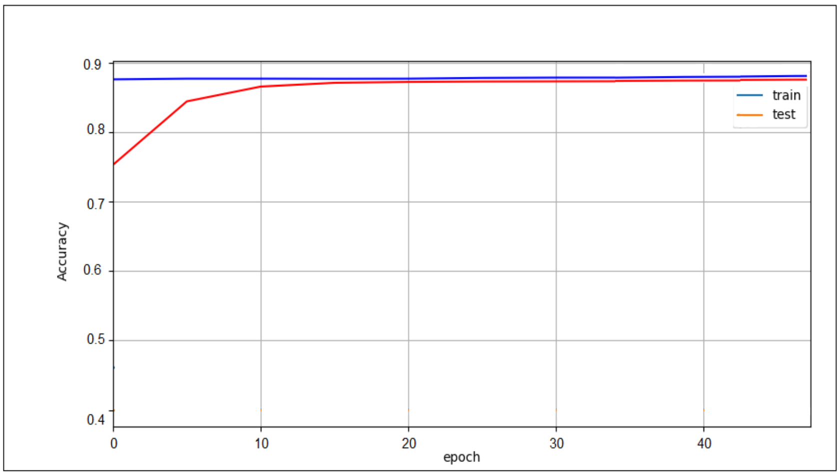

Figure 5: Trained accuracy for CB- CNN model

For the images obtained from https://www.kaggle.com/datasets/emmanuelpintelas/few-shot-datasets, the mode was trained and evaluated. There are two modes, one with and one without augmentation, and these images are used to train the model. Augmentation improves the few shot images and creates multi-view data that deep learning models can use effectively. A comparison is made between the trained accuracy of 2D CNN and the Efficient NET model, as described in [7]. With the help of an EfficientNet architecture, the authors of [7] were able to attain a 52% accuracy rate using Kaggle-obtained noisy MobileODT cervical data. Alternatively, following cleaning using the suggested Deep Cleaner method, the same architecture with ROI cropping achieved an accuracy of 76.56%. In contrast, the developed model achieves an 84.54% accuracy rate when trained on several variation images using an efficient 2D CNN model.

To get the accuracy for the 2D Convolutional Neural Network (CNN) and Efficient Net architecture of [7] in both augmented and non-augmented modes, the training data for various K samples is recorded in table 1 and graph in figure 6 below. According to the results, the 2D CNN architecture with improved layers performs better than Efficient Net, as mentioned in [7], and this is due to the fact that the proposed design efficiently handles images with multiple color channels.

Models	Accuracy in %
AE	59.5
No Cleaning	62
Manual Cleaning	67.5
DAE	73.8
Few shot Learner Efficient Net	76.56
Few shot Learner 2D CNN	84.54

Table 1: Tabulated results for Augmented and Non- Augmented images using the Efficient Net and 2D CNN model

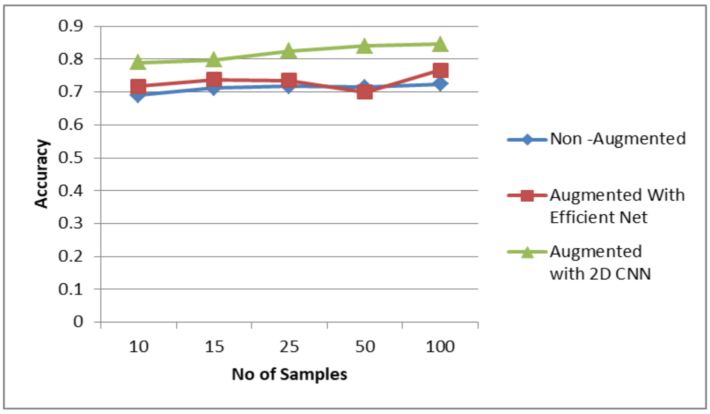

Figure 6: Performance measure of few shot images

The graph in Figure 7 compares the accuracy of a few cleaning shots with many other techniques, based on the corresponding values recorded in Table 2 when compared with that of [7].

K Samples	Non -Augmented	Augmented With Efficient Net	Augmented with 2D CNN
10	0.69	0.717	0.789
15	0.712	0.737	0.798
25	0.717	0.735	0.8253
50	0.715	0.7	0.8391
100	0.724	0.7656	0.8454

Table 2: Recorded accuracy for different model

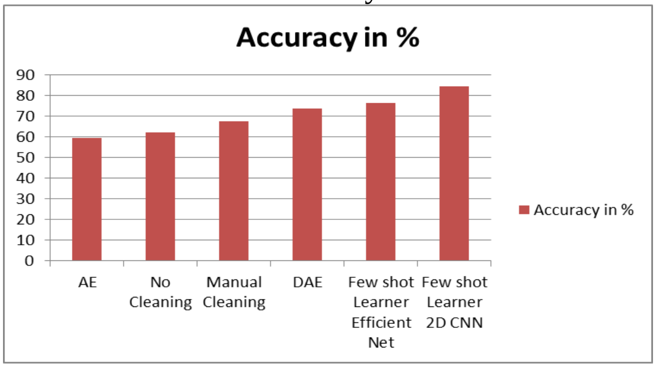

Figure 7: Accuracy Comparison

Several approaches, including those described in [7], are compared to the proposed model in order to determine how accurate it is. Auto Encoders (AE), No Cleaning, Manual Cleaning, Denoising Auto Encoders (DAE), Few Shot Learner Efficient Net, and Few Shot Learner 2D CNN were the few shot cleaning methods that were utilized for the comparison. The accuracy of the Few Shot learner 2D CNN is quite good in comparison to all of the methods that were examined; this demonstrates that the deployment of the Few Shot learning process in cleaning the dataset has improved more than the other methodologies.

4. CONCLUSION AND FUTURESCOPE

Through the utilization of the 2D-CNN and CB-CNN models, this research provided a method for hybrid cleaning of the few shot images that were obtainedfrom thepublic dataset repository like kaggle. We used a number of different datasets in order to preprocess the

images and resize them to a dimension that was previously determined. Initially, the 2D Convolutional neural network is used to these images in order to train the vast dataset of color images in order to produce the learned data. Following the completion of this procedure, the Channel boost convolution neural network is applied to the tiny dataset in order to enhance the image attributes by making use of the grayscale characteristics of the images. The trained data that was collected from the 2D CNN and the CB-CNN are then employed further to hybridize the model in order to improve the process of cleaning the few shot images. This is accomplished by deploying the model with the process of decision making and data augmentation procedures. In order to determine the accuracy of the deployed model, both without augmentation and with augmentation using alternative approaches are examined respectively. The results that were obtained make it abundantly evident that the proposed model is superior to a number of other methods, with an accuracy that is nearly 84.54%.
Both transformers and generative adversarial neural networks have the potential to be utilized in the future for the purpose of improving the proposed approach in order to get better outcomes on generic datasets.

5. REFERENCES

1. S. -H. Fan, M. -H. Lin, J. -Y. Jiang and Y. -H. Kuo, "A Few-Shot Learning Method Using Feature Reparameterization and Dual-Distance Metric Learning for Object Re-Identification," in IEEE Access, vol. 9, pp. 133650-133662, 2021, doi: 10.1109/ACCESS.2021.3116064.
2. Y. Ding and Y. Liu, "A Novel Few-Shot Action Recognition Method: Temporal Relational CrossTransformers Based on Image Difference Pyramid," in *IEEE Access*, vol. 10, pp. 94536-94544, 2022, doi: 10.1109/ACCESS.2022.3204404.
3. J. Shin, Y. Kang, S. Jung and J. Choi, "Active Instance Selection for Few-Shot Classification," in *IEEE Access*, vol. 10, pp. 133186-133195, 2022, doi: 10.1109/ACCESS.2022.3231365.
4. J. Bai *et al.*, "Class Incremental Learning With Few-Shots Based on Linear Programming for Hyperspectral Image Classification," in *IEEE Transactions on Cybernetics*, vol. 52, no. 6, pp. 5474-5485, June 2022, doi: 10.1109/TCYB.2020.3032958.
5. D. A. Chanti, V. G. Duque, M. Crouzier, A. Nordez, L. Lacourpaille and D. Mateus, "IFSS-Net: Interactive Few-Shot Siamese Network for Faster Muscle Segmentation and Propagation in Volumetric Ultrasound," in *IEEE Transactions on Medical Imaging*, vol. 40, no. 10, pp. 2615-2628, Oct. 2021, doi: 10.1109/TMI.2021.3058303.
6. J. Yao, T. Shi, X. Che, J. Yao and L. Wu, "DA-FSOD: A Novel Data Augmentation Scheme for Few-Shot Object Detection," in *IEEE Access*, vol. 11, pp. 92100-92110, 2023, doi: 10.1109/ACCESS.2023.3307490.
7. M. B. Bijoy *et al.*, "Deep Cleaner—A Few Shot Image Dataset Cleaner Using Supervised Contrastive Learning," in *IEEE Access*, vol. 11, pp. 18727-18738, 2023, doi: 10.1109/ACCESS.2023.3247500.
8. V. Clay, G. Pipa, K. -U. Kühnberger and P. König, "Development of Few-Shot Learning Capabilities in Artificial Neural Networks When Learning Through Self-Supervised Interaction," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 1, pp. 209-219, Jan. 2024, doi: 10.1109/TPAMI.2023.3323040.
9. Guangpeng wang , yongxiong wang , (member, ieee), zhiqun pan , xiaoming wang , jiapeng zhang , (Graduate Student Member, IEEE), AND JIAYUN PAN. "ViTFSL-Baseline: A Simple Baseline of Vision Transformer Network for Few-Shot Image

Classification", Received 27 December 2023, accepted 11 January 2024, date of publication 19 January 2024, date of current version 25 January 2024. Digital Object Identifier 10.1109/ACCESS.2024.3356187.

10. Q. Xu, J. Yang, H. Zhang, X. Jie and D. Bandara, "Enhancing Few-Shot Action Recognition Using Skeleton Temporal Alignment and Adversarial Training," in *IEEE Access*, vol. 12, pp. 31745-31755, 2024, doi: 10.1109/ACCESS.2024.3365448.
11. F. M. Siraj, S. T. K. Ayon, M. A. Samad, J. Uddin and K. Choi, "Few-Shot Lightweight SqueezeNet Architecture for Induction Motor Fault Diagnosis Using Limited Thermal Image Dataset," in *IEEE Access*, vol. 12, pp. 50986-50997, 2024, doi: 10.1109/ACCESS.2024.3385430.
12. Y. Huang, J. Wang and L. Wang, "Few-Shot Image and Sentence Matching via Aligned Cross-Modal Memory," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 6, pp. 2968-2983, 1 June 2022, doi: 10.1109/TPAMI.2021.3052490.
13. D. Tran-Anh *et al.*, "Integrative Few-Shot Classification and Segmentation for Landslide Detection," in *IEEE Access*, vol. 10, pp. 120200-120212, 2022, doi: 10.1109/ACCESS.2022.3220906.
14. S. -H. Park, N. S. Syazwany and S. -C. Lee, "Meta-Feature Fusion for Few-Shot Time Series Classification," in *IEEE Access*, vol. 11, pp. 41400-41414, 2023, doi: 10.1109/ACCESS.2023.3270493.
15. S. Lee, J. Lee, C. H. Bae, M. -S. Choi, R. Lee and S. Ahn, "Optimizing Prompts Using In-Context Few-Shot Learning for Text-to-Image Generative Models," in *IEEE Access*, vol. 12, pp. 2660-2673, 2024, doi: 10.1109/ACCESS.2023.3348778.
16. Ullah, I., Carrión-Ojeda, D., Escalera, S., Guyon, I. M., Huisman, M., Mohr, F., … & Vu, P. A. (2022). Meta-Album: Multi-domain Meta-Dataset for Few-Shot Image Classification. In Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track.
17. Carrión-Ojeda, D., Chen, H., El Baz, A., Escalera, S., Guan, C., Guyon, I., … & Zhu, W. (2022, December). NeurIPS'22 Cross-Domain MetaDL competition: Design and baseline results. In ECMLPKDD Workshop on Meta-Knowledge Transfer (pp. 24-37). PMLR.