iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** I **Month of publication:** January 2023

DOI: https://doi.org/10.22214/ijraset.2023.48665

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com

Muti-class Image Classification using Transfer Learning

Niharika Gupta[1], Priya Khobragade[2]
[1]*Student,* [2]*Mentor, Computer Science Department, G H Raisoni Institute of Engineering and Technology, Nagpur*

Abstract: ***Humans are very proficient at perceiving natural scenes and understanding their contents. Everyday image content across the globe is rapidly increasing and there is a need for classifying these images for further research. Scene classification is a challenging task, because in some natural scenes there will be common features in images and some images may contain half indoor and half outdoor scene features. In this project we are going to classify natural scenery in images using Artificial Intelligence. Based on the analysis of the error backpropagation algorithm, we propose an innovative training criterion of depth neural network for maximum interval minimum classification error. At the same time, the cross entropy and M3CE are analyzed and combined to obtain better results. Finally, we tested our proposed M3 CE-CEc on two deep learning standard databases, MNIST and CIFAR-10. The experimental results show that M3 CE can enhance the cross-entropy, and it is an effective supplement to the cross-entropy criterion. M3 CE-CEc has obtained good results in both databases.***
Keywords: convolution neural network, deep learning, transfer learning, pattern recognition

I. INTRODUCTION

Traditional machine learning methods mostly use shallow structures to deal with a limited number of samples and computing units.The convolution neural network (CNN) developed in recent years has been widely used in the field of image processing because it is good at dealing with image classification and recognition problems and has brought great improvement in the accuracy of many machine learning tasks.It has become a powerful and universal deep learning model.

Image classification is a task that requires a machine to be able to distinguish between different classes of objects in images. This task is challenging due to the sheer variety of objects that can appear in an image. a traditional approach to this would involve manually labelling. A new reconstruction algorithm based on convolutional neural networks is proposed by Newman et al and its advantages in speed and performance are demonstrated. Image classification is a task that requires a machine to be able to distinguish between different classes of objects in images, such as streets, buildings, seas, glaciers, forests and mountains. Traditional approaches to this task involve manually labeling each image with the categories in which the objects belong. However, as the number of categories increases, and the complexity of the images involved, manually labelling images becomes increasingly difficult and time consuming. Deep learning is an ai technique that has been shown to be effective at classifying complex images. it uses a type of artificial neural network (ANN), also known as a convolutional neural network (CNN), to extract features from each image and then train its weights and biases to recognize patterns in the data. The CNN can then be used to identify objects of different classes in the images, and make predictions about which class each object belongs to. Additionally, transfer learning can be used to further improve accuracy by training the network with data from multiple datasets, allowing for models that are more generalizable. Finally, computer vision algorithms such as edge detection, colour histogram analysis, region segmentation, and pattern recognition can also be employed to classify images

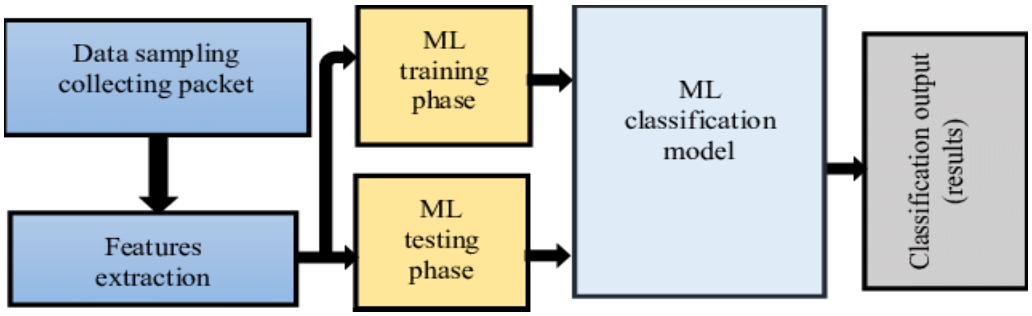
II. PROPOSED METHODOLOGY

Even though all study domains share some steps in the experimental design, the use of an ML approach must be cross-disciplinary. We can distinguish the following steps in the ML methodology used in image classification specifically: Data Collection, Data Pre-processing & Augmentation, Model Selection, Model Training, and Model evaluation and parameter tunning are the essential steps:

1) *Data Collection:* Collect data for as many classes as needed for classification. Ensure the data is labelled correctly with annotation. The data collection is followed by image annotation, the process of manually supplying details about the data's underlying reality. Image annotation, to put it simply, is the act of visually identifying the type and position of items that the AI model should be trained to recognise. In order to train the model to recognise different patterns and provide suggestions based on them, the picture or video data sets should thus include useful information. Therefore, the characteristic situations need to be captured to provide the ground truth for the ML model to learn from. For example, in industrial automation, image data needs to be collected that contains specific part defects. Therefore a camera needs to gather footage from assembly lines to provide video or photo images that can be used to create a dataset.

©IJRASET: All Rights are Reserved

2) *Data Pre-processing & Augmentation:* Perform data pre-processing such as normalization and augmentation to increase the size of the dataset. This preprocessing step would be applied to images in training and in testing. However, if the collected training data is not representative of the levels of contrast the model may see in production, there is less certainty that a constant contrast adjustment is appropriate. Instead, randomly altering image contrast during training may generalize better. This would be augmentation. If the model will be used in production on only low contrast in all situations, requiring that every image undergo a constant amount of contrast adjustment may improve model performance.
3) *Model Selection:* Select an appropriate model for the task that works well using the data input. Model selection is a key ingredient in data analysis for reliable and reproducible statistical inference or prediction, and thus central to scientific studies in the field. Model selection is the process of choosing a machine learning model that best fits the data and the goals of the project. There are many factors to consider when selecting a model, including the complexity of the model, the performance of the model on the training data, and the ability of the model to generalize to new data. Overall, the goal of model selection is to find the model that best fits the data and the project goals, while also taking into consideration factors such as model complexity and generalization ability.
4) *Training:* Train the model on the prepared dataset in order to adjust the weights and biases to better fit the data. The right response, sometimes referred to as a target or target characteristic, needs to be included in the training data. The learning algorithm generates an ML model that captures these patterns by looking for patterns in the training data that relate the qualities of the input data to the goal (the prediction you want to make). Simply said, training a model entails learning (deciding) appropriate values for each weight and bias from labelled samples. In supervised learning, an algorithm uses a large number of instances to develop a model and attempting to find a model that minimizes loss; this process is called empirical risk minimization.
5) *Evaluation:* Evaluate the performance of the model using different metrics. Model evaluation is the process of using different evaluation metrics to understand a machine learning model's performance, as well as its strengths and weaknesses. Model evaluation is important to assess the efficacy of a model during initial research phases, and it also plays a role in model monitoring.



1) *Parameter Tuning:* Tune the hyperparameters of the model to further optimize the accuracy. Hyperparameter tuning consists of finding a collection of hyperparameter variables that a learning algorithm should have in order to apply to any given data set. That hyperparameters in combination maximizes the performance of the model, minimizing a standardized loss function to produce better results with fewer errors. Hyperparameter tuning takes advantage of the processing infrastructure of Google Cloud to test different hyperparameter configurations when training your model. It can give you optimized values for hyperparameters, which maximizes your model's predictive accuracy.

The application of machine learning techniques is widespread, and more articles have been published recently in particular. These are the steps by which the model is trained and tested. So, that it can give the accepted accuracy with respect to classification.

III. DATABASE, PACKAGES

A. *Database: Intel image classification*

The Intel Image Classification dataset is a collection of images of natural scenes from around the world, organized into six categories: buildings, forests, glaciers, mountains, sea, and streets. The dataset includes approximately 25,000 images, each with a size of 150x150 pixels. The images are split into training, test, and prediction sets, with approximately 14,000, 3,000, and 7,000 images in each set, respectively.

The dataset was published by Intel on the Analytics Vidhya website as part of an image classification challenge. It can be used to train and evaluate machine learning models for tasks such as image classification and object recognition. The Intel Image Classification dataset is a collection of images of natural scenes from around the world, organized into six categories: buildings, forests, glaciers, mountains, sea, and streets. The dataset includes approximately 25,000 images, each with a size of 150x150 pixels. The images are split into three sets: training, test, and prediction. The training set contains approximately 14,000 images, the test set contains 3,000 images, and the prediction set contains 7,000 images. Intel first made the dataset available as part of an image classification challenge on the Analytics Vidhya website. It can be used to train and test machine learning models for things like object recognition and image classification. The dataset's six classes—buildings, forest, mountain, sea, and street—offer a diverse collection of images that can be used to evaluate a model's performance on a variety of natural scenes. Overall, the Intel Image Classification dataset is a valuable resource for researchers and developers working in the field of machine learning and image processing. It provides a large, diverse dataset that can be used to train and evaluate models for a variety of tasks and applications.

B. Softwares: Colab Notebook, Visual Studio

1) *Colab Notebook:* Colab is a free cloud-based software platform that allows users to develop and run machine learning models in their web browser. It is built on top of the Jupyter notebook framework, and provides an interactive environment for developing and experimenting with machine learning models. Colab is popular among machine learning researchers and developers because it allows them to quickly and easily develop and test their ideas without having to install and configure complex software on their local machines. It provides access to powerful hardware such as GPUs and TPUs, which can accelerate the training of machine learning models, and integrates with popular libraries for deep learning like TensorFlow and PyTorch Overall, Colab is a useful tool for anyone interested in machine learning and deep learning. It provides an easy-to-use platform for developing and running machine learning models, and can be accessed from any web browser.
2) *Visual Studio:* It is a software development platform developed by Microsoft. It includes a suite of tools and services for building, testing, and deploying applications, as well as tools for collaboration and source control. Visual Studio supports a variety of pro gramming languages, including C#, C++, and Python, and allows users to develop applications for a range of platforms, including Windows, macOS, and Linux. It also integrates with popular web development frameworks such as ASP.NET and Angular. Visual Studio is popular among professional software developers because it provides a rich set of features and tools that can help streamline the development process. It includes a powerful code editor, a debugger, and a variety of tools for testing and deployment, making it a comprehensive platform for building applications.

C. Packages

The packages that are used to build the model are

1) *Keras:* Keras is a Python-based open-source neural network library. It is used for building and training deep learning models, and it is designed to be user-friendly, modular, and extensible. At a high level, Keras provides a consistent set of APIs for defining and training deep learning models. One of the main advantages of Keras is its flexibility. Keras allows you to define your own custom neural network layers and operations, and it provides a wide range of pre-built layers and operations that you can use to quickly and easily build complex model architectures. It also allows you to easily switch between different backends, such as TensorFlow, Theano, and PlaidML, which makes it easy to train your models on different hardware and infrastructure. Overall, Keras is a powerful and widely-used tool for building and training deep learning models. It provides a user-friendly and modular interface that makes it easy to quickly build and experiment with different model architectures, and it is widely used by researchers and practitioners in academia and industry.
2) *TensorFlow:* Google developed this open-source software library for machine learning. It is utilized for numerous purposes, including natural language processing, computer vision, and predictive analytics.At a high level, TensorFlow allows you to define, optimize, and evaluate mathematical expressions involving multi-dimensional arrays called tensors. TensorFlow uses a system of computational graphs, where the nodes in the graph represent mathematical operations, and the edges represent the tensors that flow between them. This allows you to easily construct and modify complex machine learning models using TensorFlow's rich set of APIs and libraries. One of the main advantages of TensorFlow is its flexibility and extensibility. TensorFlow allows you to define your own custom mathematical operations, and it provides a wide range of pre-built operations that you can use to quickly and easily build machine learning models. It also provides a range of tools and libraries that make it easy to train and evaluate your models, and to deploy them in a variety of environments, such as on-premises servers, cloud platforms, and mobile devices.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

IV. ALGORITHMS

A. *Transfer Learning*

Transfer learning is a machine learning technique in which a model trained on one a model on a second, related task is used as a starting point for the first task. This allows the second model to benefit from the knowledge and experience of the first model, and can greatly speed up training and improve performance.

Transfer learning is often used when there is a lack of labeled data for the target task, or when the target task is related to the source task. For example, a model trained on image classification tasks could be used as the starting point for a model that performs object detection, since the two tasks are related. Similarly, a model trained on natural language processing tasks could be used as the starting point for a model that performs sentiment analysis, since both tasks involve working with text data.

Transfer learning is a powerful tool that can greatly speed up model training and improve performance. By leveraging the knowledge and experience of pre-trained models, transfer learning allows you to quickly and easily build and train models for new tasks, even with limited data. It is used extensively in a variety of fields, including speech recognition, natural language processing, and computer vision.

B. *Convolutional Neural Netwirks*

A type of neural network known as a convolutional neural network (CNN) is made to process data and has a grid-like structure, such as an image. It is composed of multiple layers, including pooling layers, fully-connected layers, and convolutional layers.

The convolutional layers of a CNN apply a series of filters applied to the input data that are used to extract data features. Most of the time, these filters are two-dimensional, small matrices that are applied to specific parts of the input data. By applying these filters across the entire input data, the convolutional layers are able to extract a rich set of features that capture the spatial and temporal relationships in the data. The pooling layers of a CNN are used to downsample the output of the convolutional layers, reducing the spatial dimensions of the data and increasing the robustness of the features. By reducing the number of parameters in the model and preventing overfitting, this may be of assistance. Finally, the fully-connected layers of a CNN are used to make predictions based on the extracted features. These layers typically use a softmax activation function to produce probabilities for each of the classes in the task, allowing the CNN to make multi-class predictions. Overall, CNNs are a powerful and widely-used tool for many image classification tasks, including multi-class classification. They are able to automatically learn complex patterns in the data, and can perform at the highest level on numerous tasks.

There are several types of convolutional neural networks (CNNs), which differ in the architecture and the specific operations used in the convolutional and pooling layers. Some common types of CNNs include:

1) *LeNet:* LeNet is a simple CNN architecture that was developed by Yann LeCun in the 1990s. It consists of a series of convolutional and pooling layers, followed by a few fully-connected layers. LeNet was one of the first successful applications of CNNs to practical problems, and was used for tasks such as handwritten digit recognition.
2) *AlexNet:* AlexNet is a CNN architecture that was developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton in 2012. It was the first large-scale CNN to achieve state-of-the-art performance on the ImageNet dataset, and won the 2012 ImageNet Large Scale Visual Recognition Challenge (ILSVRC). AlexNet introduced several innovations, such as the use of ReLU activation functions and dropout regularization, which are now common in many CNN architectures.
3) *VGG:* The VGG network is a CNN architecture that was developed by Karen Simonyan and Andrew Zisserman in 2014. It is characterized by its use of small, 3x3 convolution filters, which are stacked to form deep networks. VGG was one of the top performing models in the 2014 ILSVRC, and has been widely used as a starting point for many other CNN architectures.
4) *ResNet:* ResNet is a CNN architecture that was developed by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun in 2015. It is characterized by its use of skip connections, which allow the gradients in the network to flow more easily and improve the training of very deep networks. ResNet won the 2015 ILSVRC and has set new records for image classification performance on several benchmarks.
5) *Xception:* It is a 71-layer deep convolutional neural network. The ImageNet database contains a pretrained version of the network that has been trained on more than one million images. Images can be classified into one thousand object categories using the pretrained network, including keyboard, mouse, pencil, and numerous animals. As a result, the network has learned rich feature representations for a wide range of images. The network accepts images that are 299 by 299 pixels in size. The Xception model can be used to classify new images with classify. Replace GoogLeNet with Xception by following the instructions in Classify Image Using GoogLeNet. Follow the steps in Train Deep Learning Network to Classify New Images and load Xception rather than GoogLeNet to retrain the network for a new classification task.

6) VGG16 and VGG19: The concept of the VGG19 model (also VGGNet-19) is the same as the VGG16 except that it supports 19 layers. The "16" and "19" stand for the number of weight layers in the model (convolutional layers). This means that VGG19 has three more convolutional layers than VGG16. The VGG16 model achieves almost 92.7% top-5 test accuracy in ImageNet. ImageNet is a dataset consisting of more than 14 million images belonging to nearly 1000 classes. In VGG19, we use an existing model to solve different but related problems. Basically, we try to exploit what has been learned in one task and improve generalization in another task. We use the model's pre-trained weights or model architecture to solve our problem.

Overall, there are many different types of CNNs, each with its own strengths and weaknesses. The choice of architecture will depend on the specific characteristics of the data and the requirements of the task.

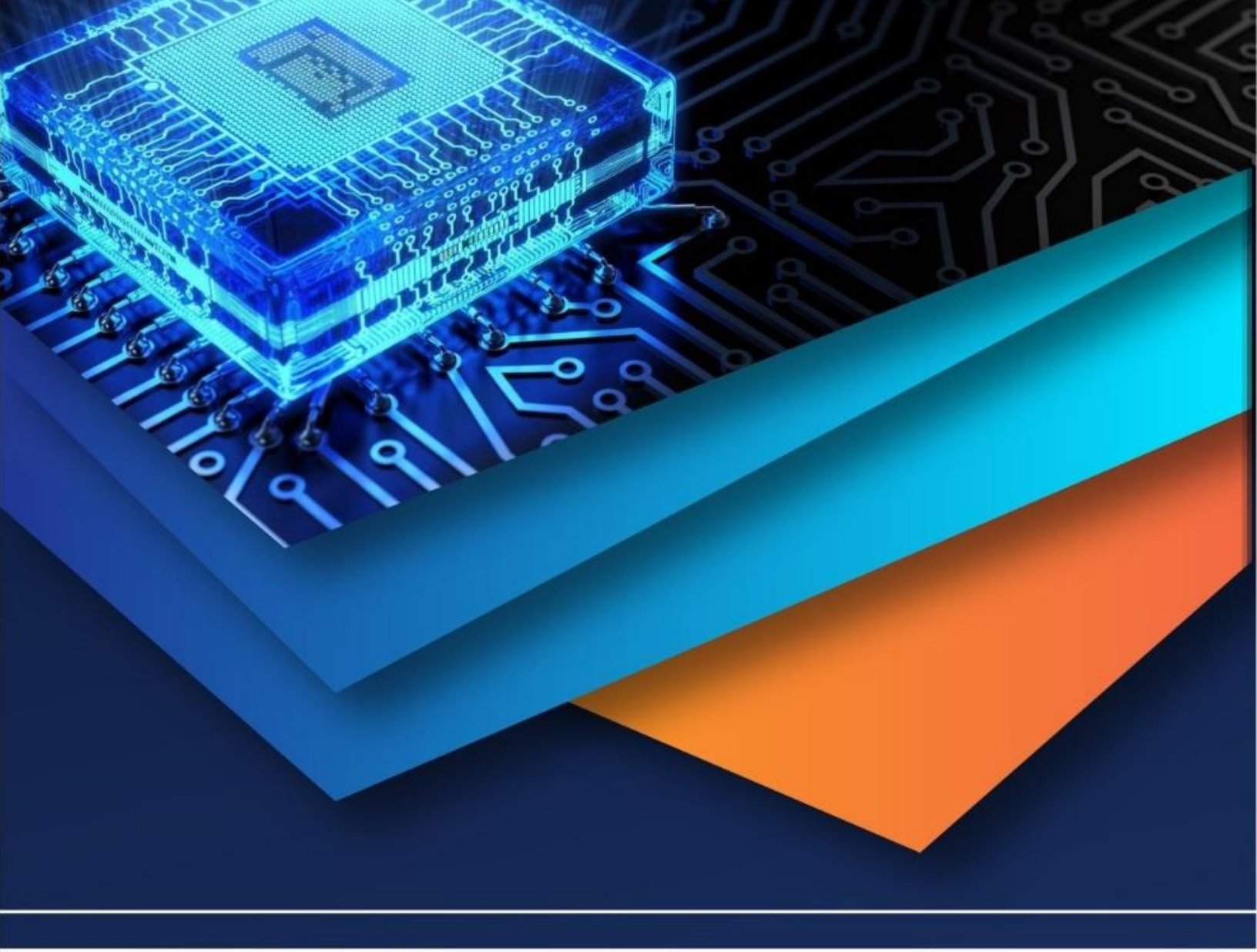
V. FUTURE SCOPE

The process of organizing data into various classes or categories based on particular characteristics is known as classification. It is a crucial step in the field of machine learning, where algorithms are trained to classify data based on previously known or labeled examples. The model, which is builded, is used to classify the images generated by various resources. We can further develop this model by increasing epochs to attain more accurate results.

The model's performance can be enhanced and its potential applications expanded by increasing the number of epochs, which is a measure of how frequently the model is exposed to the training data. Such as the recognition of objects and scenes in surveillance footage, the classification of terrain with the help of satellite imagery, the filtering and organization of personal collections of photos or videos, the assistance with the creation of artistic or creative projects, and the classification of geopolitical regions.

VI. ACKNOWLEDGEMENT

The author would like to express their appreciation to Prof. Priya Khobragade for her wise counsel and ongoing assistance during the project. They would also want to give particular thanks to Prof. Minakshee Chandankhede for her diligent supervision of the improvisation. We successfully finished this paper with your helpful guidance, and we are grateful to have both of you as our mentors.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)