International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
Article Received: 14 August 2023 Revised: 27 September 2023 Accepted: 22 October 2023

ATiTHi: A Deep Learning Approach for Tourist Destination Classification using Hybrid Parametric Optimization

Tejaswini Bhosale[1], S. Pushkar[2]
[1]Research Scholar, Birla Institute of Technology, Mesra, Ranchi-835215, India.
[2]Assistant Professor, Birla Institute of Technology, Mesra, Ranchi-835215, India.
[1,2]bhosale.tejaswini@gmail.com, shashank@bitmesra.ac.in

Abstract- A picture is best way to explore the tourist destination by visual content. The content-based image classification of tourist destinations makes it possible to understand the tourism liking by providing a more satisfactory tour. It also provides an important reference for tourist destination marketing. To enhance the competitiveness of the tourism market in India, this research proposes an innovative tourist spot identification mechanism by identifying the content of significant numbers of tourist photos using convolutional neural network (CNN) approach. It overcomes the limitations of manual approaches by recognizing visual information in photos. In this study, six thousand photos from different tourist destinations of India were identified and categorized into six major categories to form a new dataset of Indian Trajectory. This research employed Transfer learning (TF) strategies which help to obtain a good performance measure with very small dataset for image classification.VGG-16, VGG-19, MobileNetV2, InceptionV3, ResNet-50 and AlexNet CNN model with pretrained weight from ImageNet dataset was used for initialization and then an adapted classifier was used to classify tourist destination images from the newly prepared dataset. Hybrid hyperparameter optimization employ to find out hyperparameter for proposed Atithi model which lead to more efficient model in classification. To analyse and compare the performance of the models, known performance indicators were selected. As compared to the AlexNet model (0.83), MobileNetV2(0.93), VGG-19(0.918), InceptionV3(0.89), ResNet-50(0.852) the VGG16 model has performed the best in terms of accuracy (0.95). These results show the effectiveness of the current model in tourist destination image classification.

Keywords: Convolution neural network, Hyperparameter optimization, Deep learning, Tourist image classification, Transfer Learning

I. Introduction

Tourist Destination images plays an important role in the selection of appropriate destination in the tourism[1]. It refers to impression of a tourist destination as well as people's thoughts, ideas and tourist strategies. A positive tourism image promotes the destination's tourism profitability which asserts the satisfaction and loyalty of tourists, and benefits the long-term development of tourist destinations[2]. Tourism is an important socio-economic activity. India is one of the oldest civilisations in the world with a mosaic of multicultural experiences with a rich heritage and myriad of landscapes. India has a mesmeric conflation of the old and the new successions. Every year large number of tourists visit to India all across the globe[3]. Tourists are belong to various countries are interested in different orientation, culture, landscapes, pilgrim, spirituality, yoga, nature & wild life of India. Therefore choosing appropriate tourist destination plays a vital role to satisfy different orientation of tourists. Various factors such as trip motivations, personal interest, trip characteristics, destination choice, trip expenditure influence tourist decision-making process. In particular, selection of destination influence by personal interests and destination photos from social media, blogs & forums[4]. Due to advancement in technologies there is an increase in usage of internet, smart, phone, and cameras hence images data are growing exponentially, an analysis of images is a challenging research problem[5].

Content based image classification play an import role in tourist image classification. content based image analysis techniques analyses the content of image such as shapes, colours, texture are widely used for image feature extraction[6]. In content based image system every image inserted into database is analysed and stored in a form of pixel and their feature vectors. Image classification is a process of assigning a correct predefine classes to as per their extracted features. Histogram Oriented Gradient (HOG), scale-invariant feature transform (SIFT),Support Vector Machine(SVM) [7] are the traditional pattern recognition methods.

Convolution Neural network (CNN) a deep learning model used for image classification mainly are AlexNet,VGG-16,GoogleNet,ResNet etc.[8] where VGG stands for Visual Geometry Group from Oxford, AlexNet is a CNN developed by Alex Krizhevshy, GoogleNet developed by researcher at google, ResNet is a residual neural network . Deep learning

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
Article Received: 14 August 2023 Revised: 27 September 2023 Accepted: 22 October 2023

has reached the pinnacle of popularity due to its ability to learn deep representations. The capability to learn multiple levels of representations and abstractions from data makes it unique.

II. Literature Survey

For a clearer insight into the Atithi model, we studied various AI and deep learning approaches in literature. In tourism many research domains have been cover under the AI umbrella such as computer vision, deep learning, machine learning etc. Photography has been inseparable from tourism. Computer vision play important role to analyses these photographs.

2.1 AI based Destination study

In today's digital age, data posted and spread online, especially images, have substantial influences on different aspects of the intelligent configuration of destinations[9]. Further, tourists' perceptions of a destination's image after a visit affect their satisfaction and willingness to return, depending on whether the destination is able to provide a satisfactory travel experience[10]. Many researcher focuses on different study aspect of AI in tourism. Zhous acquire geospatial big data for spatial and temporal features of photos of frequently visited tourist attractions in United States to automatically identify tourist locations, which aids in the gathering of city information based on collective knowledge and local experience. They utilize the RHadoop platform to enable scalable geoprocessing workflows and to speed up geospatial problem solving[11]. S.giglio also uses geo-located images provided by Flickr to understanding relationship between human mobility and touristic attractions through ML clustering algorithm. AI based image building developed by[12]played a game-changing role in Xian city tourism marketing. Zhang model has been utilized for scene recognition, landmark recognition, and food image recognition using the transfer learning technique[10]. uses IOT enable tourist attraction system to enhance tourist experiences in Barcelona city. Carlos use IoT enable device to collect real time data then apply collaborative filtering technique & multi label classification techniques to plan different tourist activities like searching, planning etc. This help to generate Person Specific Trip but also faces many real time challenges for data collection[10]. Brazil is fast urbanizing tourist continent in south America. Better urban planning need to investigating the spatiotemporal distribution and dynamics of tourist flows in city. M Figueredo et al. apply AI techniques to make more tourist friendly mobile interface in tour design with collaborative filtering technique and fuzzy technique for scene classifier[13].Another important aspect of AI model is to differentiate between tourist and locals. Derdouri et al. model developed to understand who is tourist and who is local with help weather, mobility, and photo content data from user profiles. But scope of data collection is limited to the location information provided in the users profiles this affect the performance of model[14]. zhang et al.also study behavioral patterns between residents and tourists from user generated photos to identifying and promoting the perceived tourism destination image , improve tourist and host relationship, optimizing public infrastructure and services[13]. Some researchers integrate AI with mobile applications to assist with real-time analysis, such as locating scenic locations, taking pictures and identifying features, and planning trips. Chen uses YOLO v3 to more easily identify tourism attractions and serves as a tour guide by providing historical context for the scenic location in real-time during visitors[15].

2.2. CNN related work in destination classification

Deep learning is extensively used in various disciplines of tourism including destination classification, planning, marketing, and management[16][17] [1]. Tourism researchers use photos as data to understand how visual images represent and shape a destination. Destination photos shape and re-shape travelers' perceptions of the locations. Studies where the CNN model is applied in urban and tourism areas can be divided into two approaches: using a pre-trained model as is and using a re-trained model through transfer learning. Wang et al.(2020) used SNS to gather images of famous tourist destinations in Australia. Densnet -169 and Xception transfer learning models obtain 85.1% accuracy in destination classification of an Australian itinerary. xio et al.(2020) employs both spatial and temporal distribution characteristics of various tourism scenarios discovered through hot spot analysis. For clustering tourist scenes, xio also employs the latent Dirichlet allocation (LDA) algorithm. Xio also evaluates seasonal indexes for various tourist spots in Jiangxi, China. Y Li et al. combines deep learning model with kernel classification algorithm SVM able to identify popular landmark from Beijing, China with good accuracy but his work limited to to most popular places in Beijing[18]. CNN is used as the fundamental building component in Yong Xu's tourism scene classification with multistage transfer learning model with category hierarchical structure. In Multistage CNN achieves by constructing a baseline model and one-stage transfer learning model. It achieve new performance bounds for tourism scene classification with small number of tourist scene with predefine categorical hierarchy.Y Chang et al. apply deep learning model for classifying diverse and complex photos in urban or regional studies. Author also consider the local characteristics for various tourist destinations from Seoul using 168,216 images from 13 categories. Chang efficiently apply Inception V3 model trained on ImageNet dataset shows good performance to limited area of Seoul. J Kim et al[9]. employ addition clustering technique to automatically building categories for photo. Kim model also able to work with unstructured and uncategorized

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
***Article Received:** 14 August 2023 **Revised:** 27 September 2023 **Accepted:** 22 October 2023*

data from Isadong. Kim employ VGG-16 transfer learning model for classification of 9940 tourist destination images from 10 categories.VGG-16 model trained on places 365 dataset. Kim et al. analyzed Seoul tourism images for classifying users into residents and tourists. Kim used Flickr photos into 1000 categories using the Inception-v3 model trained on ImageNet dataset. DBSCAN applied for data clustering into predetermine categories (ref). Chen et al. uses Inner London area to understand the urban perception analytics about tourist spot with time and month using z-score techniques. ResNet-152 pertained model with Places 365 dataset utilizes to classify tourist photos to understand perception of safety, class, and uniqueness about different destinations from Inner London[19]. Zhang et al. utilizes 35,356 Flickers photo from Beijing to understand the tourist perspective about China. He employs ResNet101 deep learning model for better behavioral analysis of tourists with different cultural and geographical background[12]. P Roy et al. apply different deep learning techniques to historical destinations in Bangladesh[20].

Table 1: Summaries the destination classification work done by various researchers in world

References	Area of Research	Methodology (TF model)	Advantages	Limitations
Jiyeon Kim (2022)-10	Isadong, Korea	HDBSCAN+ VGG 16	• Automatically building categories for photo	• Categories based on data-driven manner. • Automatically clustering may leads to wrong • Only Tripadvisor is consider for data collection.
Youngok Kang (2021)-9	Seoul, Korea	Inception-v3	• Classification suitable for local characteristics. • Effectively classify a large volume of tourists photos. • Re-training a deep learning model improve accuracy.	• Results comparison is not there with other deep learning models.
R. Wang (2020)-1	Australia	CNN-Hybrid Model (Densnet -169 & Xception)	• Efficient and automated method for identifying destination images. • Automated crawler used in data collection which save time and effort. • It helps in creation of virtual destinations marketing in Australia.	• Do not discuss the potential biases in dataset. • Dataset is relatively small, consisting of only 10 destinations. • Accuracy impacted by variation in image quality • Effectiveness of proposed Model is not clear.
X Xiao (2020)-3	Jiangxi, China	DenseNet161	• Both spatial & temporal character distribution used for scenes identified using hot spot analysis • Pre-trained model is Places365 • Seasonal evaluation index is calculated • Reveals the natural characteristics and cultural heritage of the scenic spot.	• Specific social group targeted • Physical travel scenes were extracted from photos to represent travel images • Author fail to perceived emotional behind the image • Not explore to the scenic area should fully explore the tourism characteristics of different seasons.
Yuanfang Li (2020)-4	Beijing,China	CNN+SVM (Alexnet)	• Achieve good accuracy in Day images • Both view i.e. Day & night view was consider • Pretrain on Places dataset • SVM used for better prediction	• Usefulness of classifier is limited • Only 10 locations were used for classification

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
Article Received: 14 August 2023 Revised: 27 September 2023 Accepted: 22 October 2023

			• Model parameters were obtained from caffemode	
Tangquan Qi (2019)-6	China	Multistage CNN-(Alexnet)	• Multistage CNN model provide good accuracy • Category hierarchical structure used for Database creation	• Category hierarchical structure is predefined • Database contains small amount of tourism scene images.
Kun Zhang (2019)-11	Beijing, china	ResNet101	• Compare perception and behavior preference of tourists using statistic analysis and spatial analysis • Tourists behaviors and perception analysis • Time-saving and data analyzing towards a huge number of photos.	• Complex cultural and social issues not interpreted correctly • Identification of Scene outside train model is time consuming and low accuracy • Domestic Influence of flicker data • Facing problem to deal with big data.

Research Gap we identify from background review are as follows:

RG1: Frequently, we see a photo of a location that we would like to visit but are unable to determine where it is located. **A person's ability to recognize landmarks from images can be extremely useful when choosing a travel destination or identifying landmarks in a foreign country.**

RG2: No such modern research conducted in Indian Trajectory for tourist destination analysis

RG3: Different TF CNN models are not tested in tourist destination analysis. Previous researcher commonly utilizes inception v3 and Resnet for experimentation

RG4: Not uses a proper hyperparameter optimization technique for cost reduction and effective ness of CNN models

To address this research gap, our paper aims to achieve following key objectives.

Obj1: To identify different tourist destinations or travelling landmarks from India

Obj2: To prepare a large-scale tourism scene dataset for Indian tourist destinations from various categories.

Obj3: To investigate and validate hybrid hyperparameter tuning approach & data augmentation techniques for data preprocessing of destination images.

Obj4: To develop a new framework for tourism destination classification using a multi-stage transfer learning

III. Proposed Methodology

3.1. Research Process:

The analysis process of this study is shown in Figure 1. First, we extracted the posted on SNS, also collected from different local tourist agencies, friend and family members of Indian Trajectory. Second, we remove noise images not included in categories and classify photos according to the pre determine categories. Third we generate a supervise dataset of Indian tourist destinations. Fourth we split the dataset into training, testing and validation set. Fifth we applied data preprocessing techniques such as noise removal, DA on training dataset for better training of model. Later we tested various transfer learning CNN models for destination classification.VGG-16 model give relatively good performance in our experimentation. Hybrid GA-RS HPO method utilize to find best values of hyperparameters. Later we apply various data visualization techniques for performance representation like ROC curve, train-test graphs etc.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
***Article Received:** 14 August 2023 **Revised:** 27 September 2023 **Accepted:** 22 October 2023*

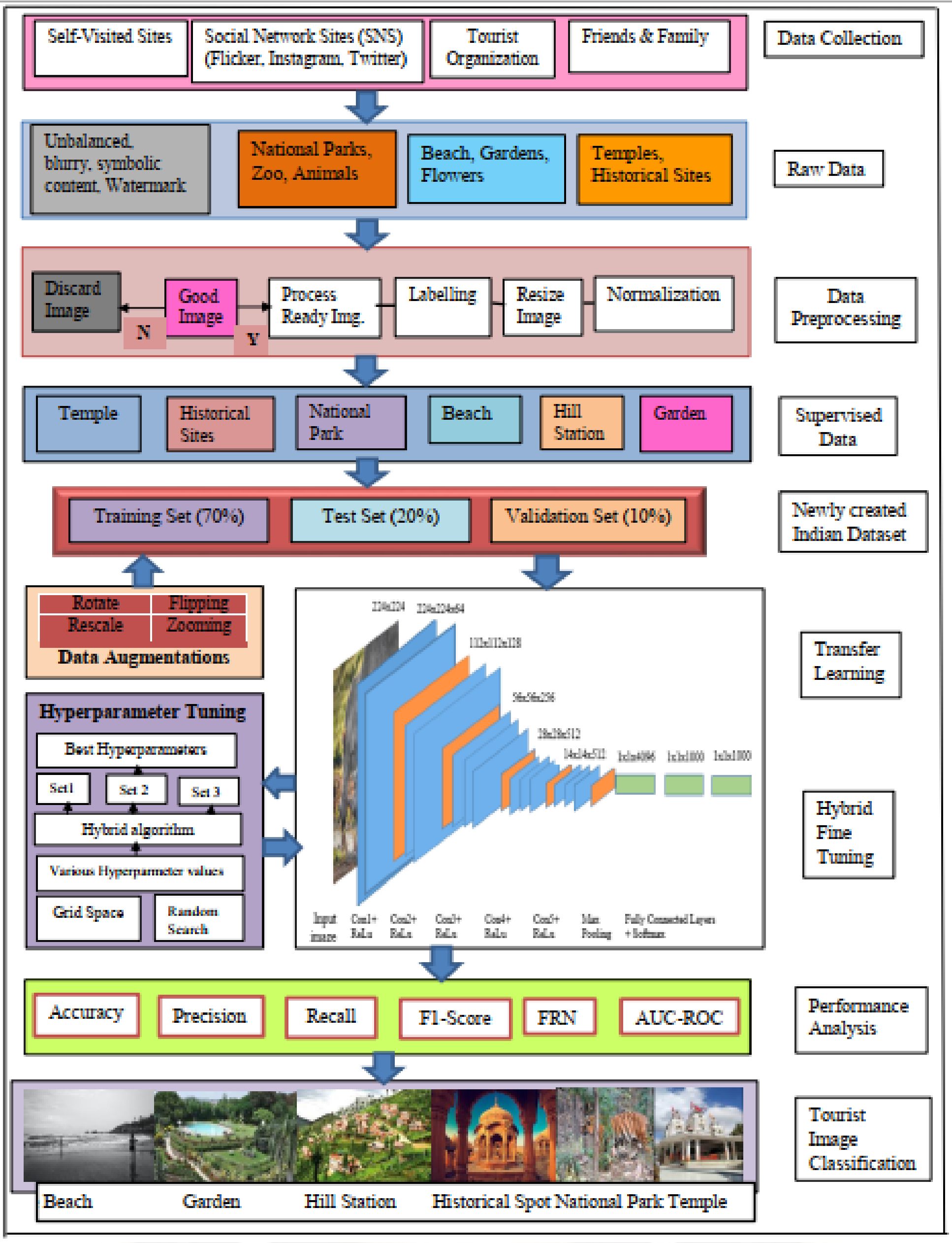

Figure 1: Research Process of Indian Tourist destination classification

3.2. Data Collection

Data collection for creation of dataset is a critical task for proposed architecture. Proposed Atithi model collected Indian tourist destination images from various social networking sites (SNS) such as flicker, Facebook, Instagram, google search etc. In recent year due to low cost internet and easy access increase large amount of data on SNS. So, SNS is popular among the researchers. Unlike other data types, SNS is updated regularly and covers a long time span and large geospace, resulting in bigger datasets with rich metadata. data offers new opportunities in tourism research by providing high spatial and temporal data that makes it possible to analyze. These platforms are also useful for planning and managing tourist's travels, to get user travel history and identify the tourists' presence in a city. this data allows researchers to highlight tourist attractions and points of interest (POIs).

3.3. Data Preprocessing:

In the realm of predictive modeling, dealing with real-world data's inconsistencies, noise, and incompleteness is crucial. The research process emphasizes various data pre-processing tasks before feeding the data into the CNN model to ensure meaningful and consistent formatting. Techniques such as data resizing and normalization are employed to create a usable dataset. All 6000 images are standardized into a square format of 250x250 pixels through cropping, maintaining central

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
***Article Received:** 14 August 2023 **Revised:** 27 September 2023 **Accepted:** 22 October 2023*

tendency. Additionally, image scaling resizes each image to 32x32 size. Data normalization, using Z-score calculation, standardizes data ranges, ensuring a uniform data distribution for efficient training. These pre-processing operations are conducted using Python 3.7.0 and verified manually. The pre-processed dataset is then randomly divided into training, validation, and testing subsets, enabling unbiased model evaluation with the help of the validation dataset. This meticulous approach ensures that the CNN model receives high-quality, consistent input data, thereby enhancing its prediction ability.

3.3.1. Data Augmentation (DA)

DA is set of algorithm that create artificial data with label-preserving transformations. Data Augmentation is one of the most useful interfaces to influence the training of Deep Neural Networks. DA help to prevent overfitting via regularization of model[21]. DA has power to express the arcane concept of semantic invariance. In Atithi model author applied different geometric image transformation for training data generation as Indian tourist destination, image dataset availability of limited data may lead to an under-fitting due to low variance in images. Tourist destination image data collection and labelling were time-consuming and costly. To lower these operational costs, we choose transforming datasets using data augmentation techniques. Atithi model utilizes various basic DA methods in geometric transform such as rotation, width_shift, height_shift, rescale, zooming, flipping. Figure 2 shows the implementation snap of DA.

```
# Set the paths to your dataset
train_data_dir = '/content/drive/MyDrive/Tourist ALL/Train'
valid_data_dir = '/content/drive/MyDrive/Tourist ALL/Valid'
test_data_dir = '/content/drive/MyDrive/Tourist ALL/Test'

# useful for getting number of output classes
folders = glob('/content/drive/MyDrive/Tourist ALL/*')

[ ] # Use the Image Data Generator to import the images from the dataset
from tensorflow.keras.preprocessing.image import ImageDataGenerator

train_datagen = ImageDataGenerator(rescale = 1./255,
                                   shear_range = 0.2,
                                   zoom_range = 0.2,
                                    rotation_range=5,
                                    width_shift_range=0.2,
                                    height_shift_range=0.2,
                                   vertical_flip=True,
                                   horizontal_flip = True,
                                   fill_mode='nearest')
```

Figure 2: Various DA operation in python 3.7.4

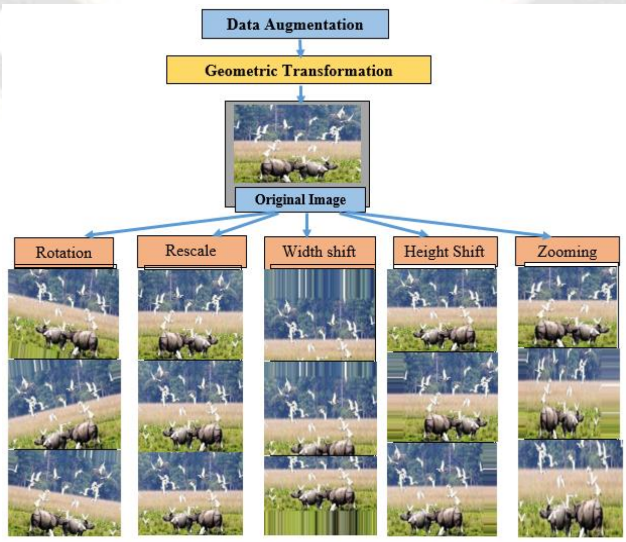

Figure 3: Geometric transformation based data augmentations effect on training dataset

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
Article Received: 14 August 2023 Revised: 27 September 2023 Accepted: 22 October 2023

Table 2 shows the data before applying DA and after applying DA. For greater accuracy and efficiency, every experimental tourist destination classification uses augmented data for training. Keras Library is used for implementation[22]. Table 2. represents proposed data models.

Table 2: Data augmentation models applied for proposed methodology

Data models	Operations	No of training images
DA 1	No data Augmentation	6000
DA 2	Geometric transform	10000

3.4. Hyperparameter optimization

Deep neural networks come with a large number of parameters or weights that are learned during training. On top of these, every neural network has additional hyperparameters that need to be configured by the user. Hyperparameter Optimization(HPO)[23] is a painstaking process to find model specific optimum values of hyperparameters. HPO is considered an NP-hard problem because it falls into the category of optimization problems that are difficult to solve efficiently. HPO, the difficulty arises from the combinatorial nature of the problem. Hyperparameters are typically discrete variables, and there is usually a large number of possible combinations. The search space grows exponentially with the number of hyperparameters, making an exhaustive search infeasible for all small problems. To find the optimal set of hyperparameters, various search algorithms and optimization techniques are employed. These include manual search, grid search, random search, Bayesian optimization. Table 3 shows the abstract view of various traditional HPO algorithms.

Table 3: Hyperparameter Optimization Methods [20]

Methods	Advantages	Disadvantages
Grid search	-It is a simple technique -can be parallelized -Reliable in low dimensional spaces -Explores the entire search space	-Sometimes trapped in local minima -suffers from the curse of dimensionality -Inefficient for high dimensional hyperparameters -Wastes time in poor performing areas -Computationally expensive
Random	-In less time and resources tries to find better solutions than grid search techniques -Can be parallelized easily for independent evaluations	-Does not guarantee an optimal result -Sometimes shows poor performance -Not reliable for training complex models
Bayesian	-Currently most promising probabilistic informed search -Mostly used for functions which are nonconvex or computationally expensive to compute	-Gausian processes possess the cubic complexity capability -Shows poor performance for high dimensional data -Limited parallelization capacity -Poor scalability

3.4.1. Hybrid Hyperparameter Optimization:

In the realm of deep learning, optimizing hyperparameters is crucial for achieving superior performance in tasks such as image classification. Hybrid HPO has innovative approach that has gained prominence, which combines the strengths of conventional techniques Grid Search and Random Search. Grid Search exhaustively explores a predefined set of hyperparameter combinations, ensuring a thorough search across the specified parameter space. On the other hand, Random Search randomly samples hyperparameter combinations, allowing for a more diverse exploration of the space[24]. By merging these techniques, proposed model can benefit from the thoroughness of Grid Search and the exploration versatility of Random Search.

So, Hybrid Hyperparameter optimization viewed as[21]:

$$\vartheta(\mathrm{h}) \approx \underset{h\in(h0,h1,h2\ldots hn)}{\arg\min} \quad L(x^{(validation)}, D_h(x^{(train)}))$$

Where D_h is a deep learning model which viewed as a functional that maps training data set X (train) to a function that minimizes a kind of loss function L(x, F),with hyperparameters are (h_i) are discrete and finite range order set such as:

$$h_i: (h_{0,}, h_1, h_2 \ldots., h_n) \qquad h_0 < h_1 < h_2 < \cdots . < h_{(n-1)} < h_n \ and\ n \in \mathbb{N}$$

This hybrid approach finds a balance between exhaustive search and computational efficiency, enabling more effective fine-tuning of hyperparameters for CNN used in tourist destination classification. As a result, this method has the potential to greatly improve accuracy and efficiency. In this work, Algorithm 1 representation of hybrid grid-random HPO.

Algorithm 1: Hybrid Grid-Random Hyperparameter Optimization

1: Set 3-D grid space for seven hyper parameters, which are to be optimized

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
Article Received: 14 August 2023 Revised: 27 September 2023 Accepted: 22 October 2023

No. of convolution & max-pool layers

No of fully connected layers

No of filters

Batch size

Activation function

No. of epoch

Learning rate

2: Identify grid boundaries

3: For I in the range of 1 to grid space boundaries

4: Perform Random sampling search

5: Evaluate samples

6: End for loop

7: Identify promising region of grid space

8: Go to step 2

9: Select the best performing model

10: Fine tune the model

11: Validate the model

Table 4 give a detail idea about the value of different hyperparameter consider for experimentation and final value get after no of iteration using Hybrid Grid Random HPO algorithm[25].

Table 4: Represent the hyperparameter values used for experimentation

Hyperparameters	Range of Value	Final Value
No of Epoch	20 to 50	50
Batch Size	32,64,120	64
Dropout rate	From 0.1 to 0.5	0.25
Learning Rate	0.0001 to 0.09	0.001
Number of layers	[1,2,4]	2
Number of Units	[8,16,32]	16
Activation Function	ReLu, Softmax	ReLu, Softmax
Deep neural network	CNN	CNN
Pooling Type	Max pooling, Average Pooling	Max pooling
Optimizer	Adam, Adamax	Adam

3.5 Convolution Neural Network:

CNN models can automatically learn high level features from raw images[26] , thus allowing for the development of applications in a shorter timeframe. CNN stack of layers used for feature extraction from images. CNN layer stack mainly consists of convolution layer, Relu, Pooling layer, dense layer with softmax activation function.

Convolution layer: A convolution is a mathematical operation applied on a matrix. This matrix is usually the image represented in the form of pixels/numbers. The convolution operation extracts the features from the image. Discrete convolution is defining as follows:

$$(p * q)(y) = \int_{-\infty}^{+\infty} p(y-t)v(t)dt \qquad (1)$$

Where *p* & *q* are two real or complex functions. They convolute into another function *(p*q)* which is normally transformed version of one of the initial functions[27].

Relu: In CNN, the Rectified Linear Units Layer integrates non-linearity and rectification in single layer which help to overcome the problem of vanishing gradient efficiently(Glorot 2011).This is activation function yields the nonlinear properties of the decision function and the overall network without affected by the cancellation problem. A rectified linear unit is a simple mathematical formulation defined as follows:

$$f\left(X_p^{(l)}\right) = \text{Max}\ (0, X_p^{(l-1)}) \qquad (2)$$

Pooling Layer: Pooling reduce the spatial size of image. Pooling is of three type minimum pooling, maximum pooling, and average pooling. Max pooling provides a form of translation invariance and thus benefits generalization[28]. It computes the maximum values over a neighbourhood in each feature map. Max pooling function is depend on the two hyperparameters, filter ($K^{(l)}$) and stride ($S^{(l)}$).

Input Size: $n_1^{(l-1)} \times n_2^{(l-1)} \times n_3^{(l-1)}$

Output Size: $n_1^{l} \times n_2^{l} \times n_3^{l}$

Max Pooling: $n_p^{(l)} = \underset{(K,S)\in n_p}{\text{Max}}\ n_{(K,S)}$

Fully connected Layer: Fully connected layer is a multilayer perceptron structure which maps higher level feature from input data. Fully Connected layer define as:

If (l-1) is a fully connected layer:

$$\mathrm{X_p^{(l)}} = \mathrm{f}\left(\mathrm{Z_p^{(l)}}\right)$$

Where: $$\mathrm{Z_p^{(l)}} = \sum_{\mathrm{q=1}}^{\mathrm{n_1^{(l-1)}}} \mathrm{w_{p,q}^{(l)}}.\mathrm{X_p^{(l-1)}} \qquad (7)$$

With goal to tune the weight parameters $\mathrm{w_{p,q}^{(l)}}$ to create a stochastic likelihood representation of predefine classes in feature map using Softmax.

3.5.1. Classification based on pre-train transfer learning model

As CNN applied widely in the research of analyzing the visual content of the photos. The conventional CNN model is not ideal because its image classification accuracy is less than 0.7[29].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
Article Received: 14 August 2023 Revised: 27 September 2023 Accepted: 22 October 2023

Theoretically, using more photos in the training process can increase a CNN model's accuracy. However, optimizing this procedure is challenging and time-consuming for specific project like tourist destination classification[30]. For this reason, a different technological technique called transfer learning was used to improve the CNN model's performance. In transfer learning (TF), the elements of a formerly trained CNN model were utilised as initialization for new CNN model.TL an effective machine learning approach which apply a CNN model that was previously trained on a large dataset and use extracted features as an initialization for a new CNN model on a much smaller dataset and redeploy it for a different reason. (Shin 2016). Even though the new classification technique is considerably different from the one for which the original model was trained, this method is useful. Representative architectures of TF- CNN include AlexNet[31], Inception V3[32], ResNet[33], VGGNet 16[8] ,VGG-19[8], MobileNetV2[34] etc.

AlexNet: AlexNet is a CNN model developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton at the University of Toronto in 2012.AlexNet work on 227x227x3 input image with 5 convolutional layers 11x11 with a stride of 4, 3 fully connected layers with 5x5 filters and stribe 2, and 2 max pooling layers uses a 3x3 window with a stride of 2. It reduces the spatial resolution of the feature maps and allows the network to learn more abstract features.

VGG-16: Very Deep Convolutional Network for Large Scale Image Recognition(VGG-16) model proposed by Karen Simonyan and Andrew Zisserman of Oxford University in 2014. VGG-16 is state of art model train using Imagenet dataset with (224* 224 * 3) input size of image followed by 13 convolutional layers with 3x3 filters, 5 max pooling layers having 2x2 windows with stribe of 2, and 3 fully connected layers with 4096 and 1000 neurons at output layer.VGG-16 is easy to train architecture with high accuracy on image classification.

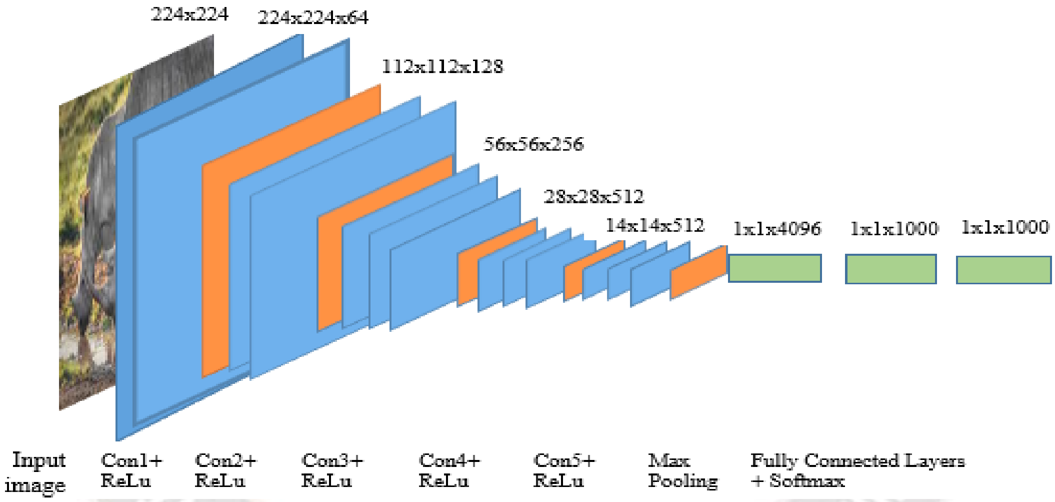

Figure 4: VGG-16 transfer learning architecture used in tourist destination classification

VGG-19: VGG-19 is another model developed by developed by Karen Simonyan and Andrew Zisserman at Oxford University in 2014. It is a 19-layer CNN that is similar to VGG-16, but with three additional convolutional layers so 16 Convolution layers in architecture. It can learn more complex feature representations than VGG-16 but computationally expensive to train and deploy.

Resnet-50: ResNet-50 is 50 layer CNN model developed by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun at Microsoft Research in 2015.Resnet-50 process 224x224x3 image size as input followed by 50 convolutional layers, including 16 residual blocks, followed by a global average pooling layer of 7x7 window with a stride of 1 and a fully connected layer with 1000 neurons. Each residual block consists of two convolutional layers and a shortcut connection. The shortcut connection allows the network to learn more complex features without increasing the number of parameters.

Inception V3: Inception v3 is a CNN model developed by Christian Szegedy et al. at Google AI in 2015 with 48-layer CNN with inception modules. Inception module allows the network to learn features at different scales. Followed by 8 max pooling layers of 2x2 window with a stride of 2, and 1 average pooling layer uses a 7x7 window with a stride of 1 which produces a single feature vector for the entire image. Inception v3 Can learn more complex feature representations of images.

MobileNetV2: MobileNetV2 is a convolutional neural network (CNN) model designed to be lightweight and efficient for mobile devices. It is based on the inverted residual block,

ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
***Article Received:** 14 August 2023 **Revised:** 27 September 2023 **Accepted:** 22 October 2023*

which consists of a bottleneck layer followed by depthwise convolutions. This structure allows MobileNetV2 to achieve high accuracy with a relatively small number of parameters and FLOPs. TF-CNN based experiment system was developed using python 3.7.0 programming with Keras & TensorFlow[22] are used here as their deep learning computing platforms. Model training is done by a cloud-based environment with GPU Tesla T4 for backend processing due to a large amount of image data.

IV. Experiment

Atithi uses CNN to retrieve features from tourist destination images to classify Indian tourist destinations using pretrained transfer learning models.TF models were trained for a total 50 epochs. The hybrid hyperparameter optimization strategy integrating grid search and random search used to select best suited hyperparameter values. Individual models and Atithi classifier model performance for each class label were evaluated using confusion matrix, ROC curve, precision, recall, F1-score, accuracy, AUC-ROC curve.

4.1 Dataset:

The collected Indian tourist destination image dataset was categorized into 6 main categories such as temples & plagiarism sites, beaches, history and heritage sites, wild life, gardens, hill Stations and each photo has been labelled with the name of the predefined category it belongs to. We used 70:20:10 ratio images for training, testing and validation of results. Table 5 shows the category wise images in dataset.

Table 5: Shows the category wise images in dataset

Categories	Training Dataset	Testing dataset	Validation dataset
Beach	7371	1727	1178
Temple	7402	1837	1182
National park	7370	1720	1169
Gardens	7375	1730	1180
Hill stations	7370	1726	1175
Heritage sites	7500	1900	1263

4.2 Experiment setup:

Experiments have been carried out on T4 GPU. The platform cast for the simulation experiments was an Intel Core i7-12700@ 32GB DDR4 RAM, and 1 TB hard disk. Python 3.8,Tensorflow,Keras ,Google colab, matplotlib, numpy, pandas are employed for conducting testing. It provides optimized runtime for DL experiments.

4.3 Performance measures:

In performance evaluation for deep learning-based classification accuracy for training data and testing, data is the total performance of a classifier represented by true predictions together with precision, recall, and F1 score. The formulation used to calculate the values of these metrics are listed as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$F1\ Score = \frac{2TP}{2TP + FP + FN}$$

Where True Positive (TP), False Positives(FP), and False Negative(FN), True Negative (TN) from confusion matrix.

V. Result Analysis

Atithi model trained using Indian Tourist destination dataset with pre-train models such as VGG-16, AlexNet, VGG-19, Inception V3,Resnet-50,MobileNetv2.Table 6 shows the architectural details like no of parameters, trainable parameters for different TF-CNN models. It also shows the training time required for each model on Indian Tourist dataset.

Table 6: Architectural details of different TF-CNN models

Model	No of Parameters	Trainable parameters	Training Time (Hrs)
VGG-16	14,865,222	150,178	2.36
MobileNetV2	3,575,878	1,317,894	3.75
VGG-19	20,174,918	150,534	2.89
ResNet-50	24,189,830	602,118	4.06
Inception V3	22,109,990	307,206	3.96
Alexnet	12,865,222	100,356	2.06

Figure 5 display the training and testing losses for TF-CNN models. Figure 5 (a) shows lowest loss nearly 0.05 for VGG-16 model while Figure 5(f) Alexnet shows major losses and inconsistent in results as compared with other models of CNN. 5(b) MobileNetv2 model display less losses like VGG-16 model while 5(c), 5(d),5(e) models VGG-19,ResNet-50,InceptionV3 respectively shows major loss during training of data model.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
***Article Received:** 14 August 2023 **Revised:** 27 September 2023 **Accepted:** 22 October 2023*

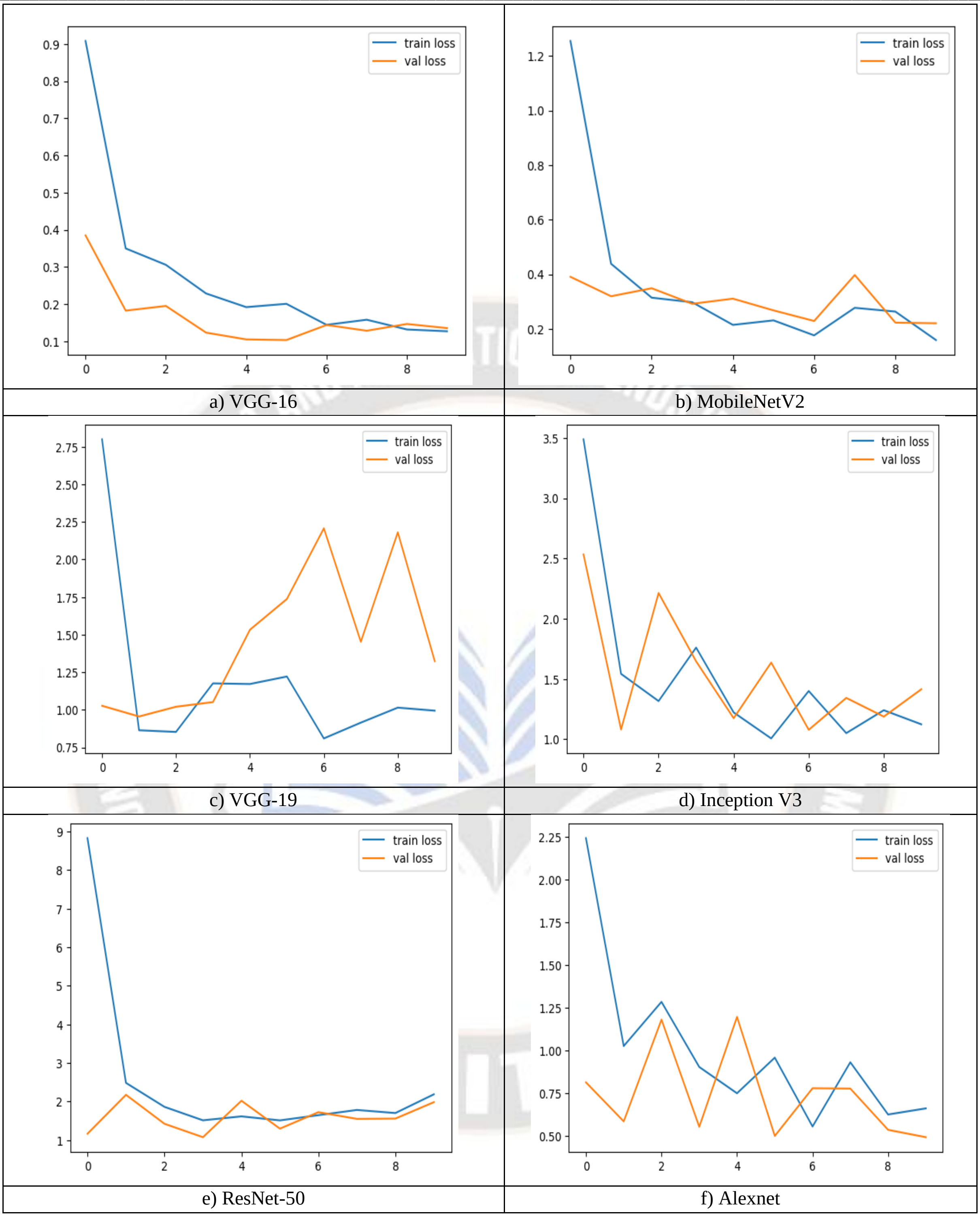

Figure 5: Training & validation loss graphs for all CNN-TF models for Tourist destination classification

Figure 6 represent various accuracy graphs for training and validation dataset. 6(a) show approximately 95.00% accuracy on validation dataset using VGG-16 model. 6(b) represent train and validation accuracy for mobilenetv2 model which is approximately 93.78%.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
***Article Received:** 14 August 2023 **Revised:** 27 September 2023 **Accepted:** 22 October 2023*

Figure 6: Training & validation accuracy graphs for all CNN-TF models for Tourist destination classification

Figure 7 display performance analysis of multiclass classification model Atithi using Area Under the Curve (AUC) and Receiver Operating Characteristic (ROC) curves. AUC and ROC are essential techniques for assessing how well categorization models work. The trade-off between a model's true positive rate and false positive rate across different threshold settings is depicted by the ROC curve. It offers a thorough comprehension of a model's capacity for class discrimination. These tools are crucial in result analysis as they offer a concise and visual representation of a model's predictive ability, allowing for comparative assessments between different models and aiding in selecting the most appropriate model for a specific task.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
Article Received: 14 August 2023 Revised: 27 September 2023 Accepted: 22 October 2023

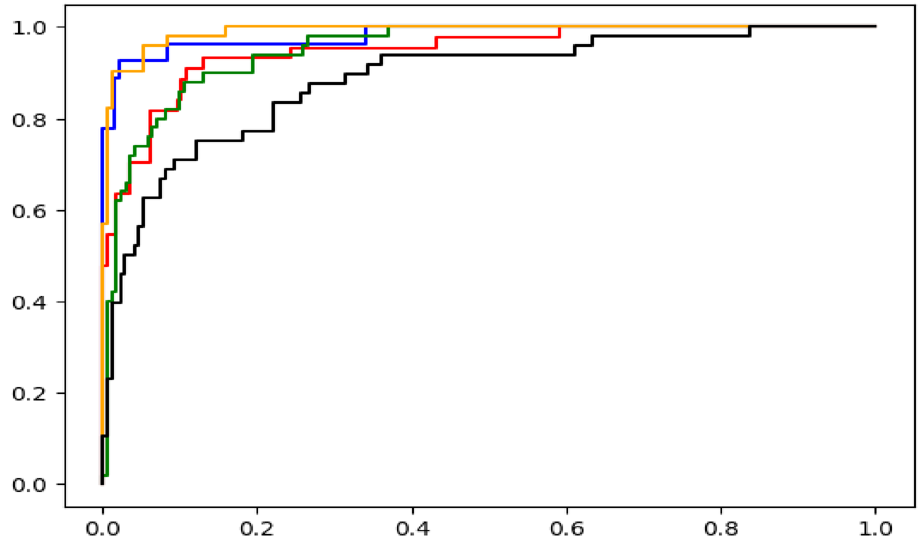

Figure 7: ROC –AUC curve of Multiclass tourist destination model-Atithi

Table 7: CNN-TF model validation performance assessment based on each class label for tourist destination classification

Model	Metrics (%)	Beach	Garden	Hill Station	National Park	Temple	Historical Places
VGG-16	Accuracy	95	93	95.22	94.44	93	90.22
	Precision	96	94	95.03	89.8	92	80.16
	Recall	89	95	85	89	69.51	81
	F1 Score	92	96	86.96	89.44	81	70.59
	AUC	94.93	93.25	91.56	90.77	89.2	78.26
MobileNetV2	Accuracy	94	92.89	91.37	90.56	89	82
	Precision	95	93	90.2	89	87	80
	Recall	91.2	92	91.26	88	86.06	78.26
	F1 Score	91.86	94.02	92	91.02	84.29	78.87
	AUC	94.03	93.81	89.96	90.63	86.78	77.2
VGG-19	Accuracy	91.8	89.03	88	90.12	77	78.33
	Precision	92.05	91	89.23	91.43	79.24	73.21
	Recall	89.65	89	91	92	88	78
	F1 Score	90	91.03	93	91	87.26	84.21
	AUC	90	87.7	89	88	84.11	83.36
InceptionV3	Accuracy	89	88.7	84.23	85.97	79.01	78.45
	Precision	91.03	89.27	85.07	86.04	80.02	79.65
	Recall	89.57	87.3	84.09	84.12	78.21	77.36
	F1 Score	89.2	89.12	84.12	83.23	79.23	78
	AUC	91.21	90.06	87	89.1	81.26	78.22
ResNet-50	Accuracy	85.2	83	82	79	73.18	70.89
	Precision	87	84	84.27	77.97	73.28	72.77
	Recall	85	82.67	83	80.27	71.39	68.97
	F1 Score	83.45	82	82.67	79.26	70.04	69.27
	AUC	85	81.96	83	80	74.2	70.89
AlexNet	Accuracy	83	82.96	83.4	79	68	66.77
	Precision	82	81.23	82.45	77	66	67.27
	Recall	84.12	80.26	81.26	78.01	67.28	64.2
	F1 Score	83.67	81.01	80	78.27	68	65
	AUC	81.2	77.96	80.66	79.82	75.81	71.26

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
Article Received: 14 August 2023 Revised: 27 September 2023 Accepted: 22 October 2023

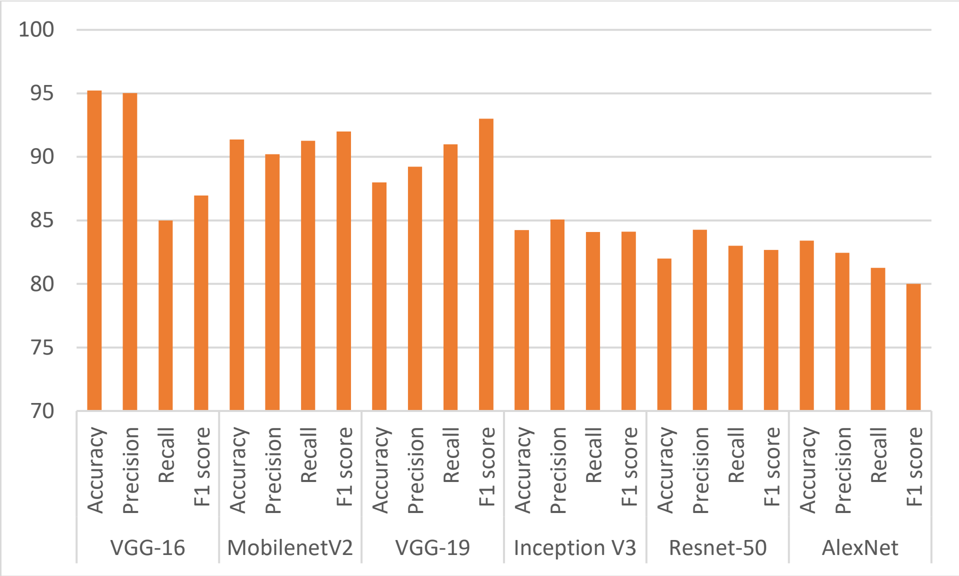

Figure 8: Comparative analysis of different TF-CNN models for tourist destination multiclass classification models

Figure 8 represent the bar-chart comparative analysis of different TF-CNN model on parameters like accuracy, precision, recall, F1-score.VGG-16 model perform able to achieve high accuracy as compare with all others. But recall of VGG-16 model is less than mobileNetV2 and VGG-19. The precision of VGG-16, VGG-19 compactible with each other. AlexNet model shows the low performance as in multiclass classification.

VI. Conclusion

In this paper, thousands of tourism destination photos, scenes of tourist destinations were extracted and quantified using CNN based image classification approach and classified into six predefined categories. The pretrained weights of the ImageNet dataset with the help of the TL technique were used for initialization of VGG-16, VGG-19, MobileNetV2, InceptionV3, ResNet-50 and AlexNet models to develop a new Atithi model for the classification of tourist destination images by utilizing a hybrid hyperparameter optimization technique The newly developed Atithi model for tourist destination classification consists of two dense layers, one with 500 channels, 0.25 dropout retention rate and a batch normalization operation. Another dense layer with 6 channels and softmax function was used to classify images into predefine six categories. Moreover, the hybrid hyperparameter optimization strategy integrating grid search and random search allowed for an effective exploration of the hyperparameter space, enhancing the model's performance by identifying optimal parameter configurations. The findings revealed that the combination of transfer learning and hybrid hyperparameter optimization can significantly increased the accuracy and robustness of the CNN model in classifying tourist destination images. Furthermore, the outcomes demonstrated that this methodology is not only efficient in its computational requirements but also provides a scalable and adaptable framework for various tourist destination image classification tasks. The best accuracy was noted in VGG16 model (0.95) as compared with VGG-19, MobileNetV2, InceptionV3, ResNet-50 and AlexNet models in Indian tourist trajectory. These results show the better effectiveness of the pretrained VGG-16 model in tourist destination image classification. In conclusion, the amalgamation of transfer learning based VGG-16 model with hybrid hyperparameter optimization of CNNs presents a robust framework for classifying tourist destination images help for better and satisfying tour. It offers a foundation for further innovation in the field of the tourism industry in India and contributes to the growing field of computer vision.

Despite the promising results, there is still room for further exploration and improvement. Research in the future could involve experimenting with different CNN architectures, exploring additional hyperparameter optimization methods, and expanding the dataset to cover a wider variety of tourist destinations.

References-

[1] X. Xiao, C. Fang, and H. Lin, “Characterizing tourism destination image using photos’ visual content,” ISPRS Int. J. Geo-Information, vol. 9, no. 12, 2020, doi: 10.3390/ijgi9120730.

[2] J. L. Crompton, “An Assessment of the Image of Mexico as a Vacation Destination and the Influence of Geographical Location Upon That Image,” J. Travel Res., vol. 17, no. 4, pp. 18–23, 1979, doi: 10.1177/004728757901700404.

[3] T. Bhosale and S. Pushkar, “Tourist spot classification using convolution neural network,” SSRN Electron. J., 2021, doi:

10.2139/ssrn.3884676.

[4] H. Kim and S. Stepchenkova, "Effect of tourist photographs on attitudes towards destination: Manifest and latent content," Tour. Manag., vol. 49, pp. 29–41, 2015, doi: 10.1016/j.tourman.2015.02.004.

[5] J. Li, L. Xu, L. Tang, S. Wang, and L. Li, "Big data in tourism research: A literature review," Tour. Manag., vol. 68, pp. 301–323, 2018, doi: 10.1016/j.tourman.2018.03.009.

[6] H. Li, J. Wang, M. Tang, and X. Li, "Polarization-dependent effects of an Airy beam due to the spin-orbit coupling," J. Opt. Soc. Am. A Opt. Image Sci. Vis., vol. 34, no. 7, pp. 1114–1118, 2017, doi: 10.1002/ecs2.1832.

[7] N. D. Hoang and V. D. Tran, "Image Processing-Based Detection of Pipe Corrosion Using Texture Analysis and Metaheuristic-Optimized Machine Learning Approach," Comput. Intell. Neurosci., vol. 2019, 2019, doi: 10.1155/2019/8097213.

[8] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., pp. 1–14, 2015.

[9] D. Kim, Y. Kang, Y. Park, N. Kim, and J. Lee, "Understanding tourists' urban images with geotagged photos using convolutional neural networks," Spat. Inf. Res., vol. 28, no. 2, pp. 241–255, 2020, doi: 10.1007/s41324-019-00285-x.

[10] J. C. Cepeda-Pacheco and M. C. Domingo, "Deep learning and Internet of Things for tourist attraction recommendations in smart cities," Neural Comput. Appl., vol. 34, no. 10, pp. 7691–7709, 2022, doi: 10.1007/s00521-021-06872-0.

[11] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, "Places: A 10 Million Image Database for Scene Recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 6, pp. 1452–1464, 2018, doi: 10.1109/TPAMI.2017.2723009.

[12] F. Sheng, Y. Zhang, C. Shi, M. Qiu, and S. Yao, "Xi'an tourism destination image analysis via deep learning," J. Ambient Intell. Humaniz. Comput., vol. 13, no. 11, pp. 5093–5102, 2022, doi: 10.1007/s12652-020-02344-w.

[13] M. Figueredo et al., "From photos to travel itinerary: A tourism recommender system for smart tourism destination," Proc. - IEEE 4th Int. Conf. Big Data Comput. Serv. Appl. BigDataService 2018, pp. 85–92, 2018, doi: 10.1109/BigDataService.2018.00021.

[14] A. Derdouri and T. Osaragi, A machine learning-based approach for classifying tourists and locals using geotagged photos: the case of Tokyo, vol. 23, no. 4. Springer Berlin Heidelberg, 2021.

[15] Y. C. Chen, K. M. Yu, T. H. Kao, and H. L. Hsieh, "Deep learning based real-time tourist spots detection and recognition mechanism," Sci. Prog., vol. 104, no. 3_suppl, pp. 1–19, 2021, doi: 10.1177/00368504211044228.

[16] R. Wang, J. Luo, and S. (Sam) Huang, "Developing an artificial intelligence framework for online destination image photos identification," J. Destin. Mark. Manag., vol. 18, no. August, p. 100512, 2020, doi: 10.1016/j.jdmm.2020.100512.

[17] Y. Li, X. Li, and C. Yue, "Recognition of Tourist Attractions," pp. 1–5, 2016.

[18] K. Zhang, Y. Chen, and C. Li, "Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of Beijing," Tour. Manag., vol. 75, no. May, pp. 595–608, 2019, doi: 10.1016/j.tourman.2019.07.002.

[19] M. Chen, D. Arribas-Bel, and A. Singleton, "Quantifying the characteristics of the local urban environment through geotagged flickr photographs and image recognition," ISPRS Int. J. Geo-Information, vol. 9, no. 4, 2020, doi: 10.3390/ijgi9040264.

[20] P. Roy, J. H. Setu, A. N. Binti, F. Y. Koly, and N. Jahan, "Tourist Spot Recognition Using Machine Learning Algorithms," Lect. Notes Data Eng. Commun. Technol., vol. 131, no. January, pp. 99–110, 2023, doi: 10.1007/978-981-19-1844-5_9.

[21] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," J. Big Data, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0197-0.

[22] M. Abadi et al., "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," 2016, [Online]. Available: http://arxiv.org/abs/1603.04467.

[23] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," J. Mach. Learn. Res., vol. 13, pp. 281–305, 2012.

[24] H. Jiang and E. Learned-Miller, "Face Detection with the Faster R-CNN," Proc. - 12th IEEE Int. Conf. Autom. Face Gesture Recognition, FG 2017 - 1st Int. Work. Adapt. Shot Learn. Gesture Underst. Prod. ASL4GUP 2017, Biometrics Wild, Bwild 2017, Heterog. Face Recognition, HFR 2017, Jt. Chall. Domin. Complement. Emot. Recognit. Using Micro Emot. Featur. Head-Pose Estim. DCER HPE 2017 3rd Facial Expr. Recognit. Anal. Challenge, FERA 2017, pp. 650–657, 2017, doi: 10.1109/FG.2017.82.

[25] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., pp. 1–15, 2015.

[26] P. R. Chandre, P. N. Mahalle, and G. R. Shinde, "Machine Learning Based Novel Approach for Intrusion Detection and Prevention System: A Tool Based Verification," in 2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN), Nov. 2018, pp. 135–140, doi: 10.1109/GCWCN.2018.8668618.

[27] P. R. Chandre, "Intrusion Prevention Framework for WSN using Deep CNN," vol. 12, no. 6, pp. 3567–3572, 2021.

[28] P. Chandre, P. Mahalle, and G. Shinde, "Intrusion prevention system using convolutional neural network for wireless sensor network," IAES Int. J. Artif. Intell., vol. 11, no. 2, pp. 504–515, 2022, doi: 10.11591/ijai.v11.i2.pp504-515.

[29] R. Kumari, A. Nigam, and S. Pushkar, "Machine learning technique for early detection of Alzheimer's disease," Microsyst. Technol., vol. 26, no. 12, pp. 3935–3944, 2020, doi: 10.1007/s00542-020-04888-5.

[30] J. E. Luján-García, C. Yáñez-Márquez, Y. Villuendas-Rey, and O. Camacho-Nieto, "A transfer learning method for

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8548
***Article Received:** 14 August 2023 **Revised:** 27 September 2023 **Accepted:** 22 October 2023*

pneumonia classification and visualization," Appl. Sci., vol. 10, no. 8, 2020, doi: 10.3390/APP10082908.

[31] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, 2015, doi: 10.1007/s11263-015-0816-y.

[32] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 2818–2826, 2016, doi: 10.1109/CVPR.2016.308.

[33] S. Targ, D. Almeida, and K. Lyman, "Resnet in Resnet: Generalizing Residual Architectures," pp. 1–7, 2016, [Online]. Available: http://arxiv.org/abs/1603.08029.

[34] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.