Climate Classification Using Geo-Tagged Images

Prof. P.V.Barekar
Department of Computer Technology
YCCE
Nagpur, India

Dnyanada Argulwar
Department of Computer Technology
YCCE
Nagpur, India

Akanksha Nagdeote
Department of Computer Technology
YCCE
Nagpur, India

Radhika Charmode
Department of Computer Technology
YCCE
Nagpur, India

Samruddhi Upganlawar
Department of Computer Technology
YCCE
Nagpur, India

***Abstract*--Given a photo of a landscape, we trained models to predict the Koppen & Geiger climate type depicted in the image. While similar studies have attempted to use satellite imagery for this, few have directly used ground level photos naturally taken from around the world especially from specific region of Indian territories. To accomplish this task, we built a dataset of photos of varying climate region and then trained SVM, logistic regression model and transferred learning convolutional neural network. Similarly, we are taking Koppen classification model and creating own hybrid classification to maintain the accuracy for individual photos to check the climate region of the location of photos where it's taken from by collecting the GPS location data for accurate classification of particular region whether GPS data usage were never the part of traditional classification model, we are maximizing the accuracy of the climate classification by making exaggerate changes in the model by not only using individual photo given but also the entire dataset of photos to classify on the map using convolution neural network and deep learning techniques and GUI technologies to showcase the effects and results.**

I. INTRODUCTION

There has been a rise in interest within the public for the system close to humans. Often, it's believed that a deeper understanding of the ecological context one lives in makes them additional doubtless to advocate for environmental policy, and because of this, it's vital to cultivate this accrued interest. In spite of this, it's typically tough to tell the additional concerns of the system around a location while not being skilled, particularly when one doesn't have the precise vocabulary to explain the system. A way to ameliorate this is often by providing the climate classification for a selected image. Though climate isn't the sole variable that characterizes associate degree system, it's a straight-forward piece of data that has a jumping off purpose to be told for the additional concerns of the nature normally. What is more, with the increase of world warming, climate is a vital facet of the system, as completely different climates could modify drastically within the future. Finally, climate may be a sensible alternative for public education, as systems like the Koppen & Hans Geiger climate classifications square measure are well-structured and outlined. We wish to make a tool to acknowledge climate class supported photos. With this, one may take an image of the character they will be around, and like a shot understand in what system they're actively moving. Rather than associate degree abstract plan, the system can then be tangible. What is more, this might function as a general instructional tool, permitting folks to be additional at home with pictures of ecosystems they will stumble across. The classifier takes in a picture of a natural landscape as input and outputs the climate that's possibly shown within the picture. During this work, we have a tendency to use many completely different machine learning paradigms to try to attain this aim, with best performance employing a (Resnet) Convolutional Neural Network.

II. LITERATURE SURVEY

1. MACHILE LEARNING ALGORITHMS LIKE RANDOM FOREST ARTIFICIAL NEURAL NETWORK AND SUPPORT VECTOR ARE COMPARED FOR SUPERVISED CROP TYPE CLASSIFICATION

Some of the most important applications of remote sensing are the classification and recognition of agricultural crop types. Till date, studies have only compared the performance and usability of the few machine learning algorithms that have emerged in the last years. Taking this in account, three different state-of-the-art machine learning classifiers namely Support Vector Machine (SVM), Artificial Neural Network (ANN) and Random Forest (RF) are compared. In the light of this purpose, a dataset of more than 500 crop fields situated in the Canadian Prairies is classified with a stratified randomized sampling approach. The mean overall classification accuracies and the standard deviations were compared. Analysis of the classification accuracy of single crops was done. SVM showed more accurate results as compared to ANN and RF. [1]

2. VECTOR MACHINES FOR FOREST MAPPING USING DECISION TREE

Automatic tree species classification on single tree level for a large area is performed by developing an automatic supervised classification strategy and finding the necessary data sources. For the estimation of additional forest parameters like diameter at breast height and volume, the derived forest map is used in a virtual forest test-bed at single tree and stand level. The calculated data is used to populate the virtual forest database which in turn can be used for the implementation of future forest development. This can be achieved by a support vector machine based decision tree. Additional improvements such as the use of LIDAR height and intensity data and comparison of SPOT and RapidEye results can be added.[2]

3. STUDY OF VARIOUS MACHINE LEARNING ALGORITHMS USED FOR THE CLASSIFICATION IN SEMIARID WOODLAND USING RAPIDEYE IMAGES

As a result of the change in leaf structure and orientation due soil moisture constraints, classification of different tree species in semiarid areas can be challenging. Machine Learning algorithms are used for the classification of 5 tree species in mopane woodland of Botswana. Limited training examples are used. Random Forest (RF) and Support Vector Machine (SVM) were used for classification. The accuracy in case of SVM was 88.75% and in RF was 85%. Demonstration of the new red-edge band in the Rapid-Eye sensor showed that it has the ability to classify tree species in semiarid environments when integrated with other standard bands.[3]

4. BASED ON MULTIPLE EARTH OBSERVATION DATA CLASSIFICATION OF LOCAL CLIMATE ZONES ARE STUDIED

This was specifically focused on urban climate evaluated SVMs, RFs, and Neural Network to identify urban climates from satellite imagery. Accuracy of 97.4% and 95.3% was achieved on the neural network and RF respectively. Recently, considerable progress was made in the determination of urban morphologies from different earth observation datasets. A relevant field of application for such methods is urban climatology since specific urban morphologies produce distinct micro-climates. However, application and comparability are so far limited by the variety of technologies used for the description of urban surfaces in earth observation. In this, local climate zones are studied.[4]

5. ECOTOPE MAPPING USING AIRBORNE HYPERSPECTRAL IMAGERY ON RANDOM FOREST AND ADABOOST TREE-BASED ENSEMBLE CLASSIFICATION AND SPECTRAL BAND SELECTION IS EVALUATED

Environmental evaluation demands the detailed land cover/land use classification at ecotope level. The possibility of the usage of airborne hyperspectral imagery for the classification of ecotopes is investigated. Based on standard classification accuracy, training time and classification stability are used to assess two tree-based ensemble classification algorithms: Random Forests and Adaboost. However, the results show that Adaboost and Random Forests show almost same accuracy of upto 70%. The neural network classifier has an accuracy of 63.7%.[5]

6. CLASSIFICATION OF TREE SPECIES ON RANDOM FOREST

This project focuses on the classification of trees species based on the Random forests (RF).This classification is a new and most powerful statistical classifier that is well established in other fields but is lesser known in ecology. Ecological data are high dimensional with complex interactions among variables, and with number of missing values among measured variables. In recent years, classification of trees is in great use by ecologists because of their easy interpretation, high classification accuracy, and ability to differentiate complex interactions among variables. Random Forests is one such method. RF is already widely used in bioinformatics, but has not yet utilized by ecologists. As the name suggests, RF combine many classification of trees to produce accurate classifications. By-products of the RF calculations include measures of variable importance and measures of similarity between data points that may be used by clustering, multidimensional scaling, graphical representation, and missing value imputation. RF to ecology include regression, multidimensional scaling ,survival analysis, ,clustering, and detecting general multivariate structure through unsupervised learning, missing value imputation and classification.[6]

7. HUMAN DETECTION IS DONE USING HISTOGRAM OF ORIENTED GRADIENT

This project is based on robust visual beholding; adopting linear SVM based human detection as a test case .Detecting citizenry in images could also be a most challenging task thanks to their variable appearance and the wide ranges of poses that they're going to adopt. The first need could even be a strong feature set that allows the clean distinction of human form. After the review of the prevailing edge in image and gradient based descriptors, it can be seen that grids of Histograms of Oriented Gradient (HOG) descriptors significantly outperform existing sets of features for human detection. Study the influence of every stage of the computation on performance, concludes that fine-scale gradients, fine orientation binning, relatively coarse spatial binning, and high-quality local contrast normalization in overlapping descriptor blocks are all important permanently results.. We study the issue of feature sets for human detection, showing that locally normalized Histogram of Oriented Gradient (HOG) descriptors provide excellent performance.[10]

8. IMAGE RECOGNITION USING DEEP RESIDUAL LEARNING IS EVALUATED

Deeper neural networks are most difficult to train. Deep networks naturally integrate low/mid/high level features and classifiers present in an end-to-end multilayer fashion, and therefore the “levels” of features are often enriched by the amount of stacked layers (depth). They explicitly reformulate layers as learning residual functions with regard to a layer inputs, rather than learning unreferenced functions. They provide comprehensive empirical evidence showing that these residual networks are easier to compress and may gain more accuracy from considerably increased depth. The depth of representations is of central importance for several visual recognition tasks.[9]

9. MACHINE LEARNING ALGORITHMS LIKE PIXEL-BASED AND OBJECT BASED IMAGE ANALYSIS FOR THE CLASSFICATION OF AGRICULTURAL LANDSCAPES USING SPOT-5 HRG IMAGERY

Pixel-based and object-based image analysis approaches are compared by use of three supervised machine learning algorithms: decision tree (DT), random forest (RF), and the support vector machine (SVM).The classification of land use and land cover (LULC) from a remotely sensed imagery can divided into the two general image analysis approaches: i) classifications based on pixels, and ii) classifications based on objects. While pixel-based analysis approach has long been the main approach for classifying remotely sensed imagery, object-based image analysis have been increasingly commonplace over the last decade .In this study, pixel-based classifications utilized few variables, achieved more similar classification accuracies, and required less time to produce than object-based classifications. Overall accuracy reports, there is no advantage to preferring one image analysis approach over another for the purposes of the mapping land cover types in agricultural environments using medium spatial resolution earth observation imagery.[7]

10. TRAINING ON MIXED SPECTRAL RESPONSES FOR CLASSIFICATION BY A SVM SMALL TRANING DATA SET WITH MIXED PIXELS IMAGES ARE USED

In this project, accuracy of a supervised image classification may be a function of the training data used generation. It is critical that the training stage of a supervised classification is design to provide the necessary information. Guidance on the planning of the training stage of a classification of typically involves the utilization of an outsized sample of the randomly selected pure pixels so as to characterize the different classes. Such guidelines are generally made without regard to the specific nature of the application in-hand, including the classifier to be used. The design of the training stage would really be based on the classifier to be used since individual training cases can vary in values as can any training set to a range of classifiers. It is argued that the training stages can design on the basis of the way the classifier operates and with emphasis on desire to separate the different classes rather than describe each of them. This approach to training of a support vector machine (SVM) classifier that's the other of that generally promoted for training set design was suggested. This approach use a little sample of mixed spectral responses which drawn from purposefully selected different locations (geographical boundaries) in training. The approach is based on mixed pixels which were normally masked-out of analyze as undesirable and problematic. A sample of such data should, however, be easier and cheaper to acquire than that suggest by conventional approach.[8]

11. ADVANCE CLASSIFICATION ON THE IMAGES OF FOREST IN NEW YORK

The project focuses on the classification of images for the area around the Heiberg Memorial Forest in Tully, New York. This requires manipulation and organization of existing forest. This project aims to use different classification methodologies. Traditional approaches such as supervised classification. It provides baseline classifications of satellite imagery (Landsat) and also focus on high spatial resolution of images. Such fundamental techniques provide a structured foundation so as to make the comparisons with other analyses. The main focus of this paper is generating species level classification from Landsat Enhanced Thematic Mapper Plus (ETM+) imagery. This analysis classifies the Landsat imagery collected during spring, summer, and fall seasons. The project will consider algorithms for topographic analyses. Lidar-derived data layers is limited which is been documented by utility of incorporating which is more challenging. Such analysis required alternative methods such as rule-based classifiers or neural networks. Both the supervised classifier and the rule-based approach provides reference information for classification and assessment. The results attained for the classification that it is possible to develop reasonable overall accuracy while performing a species level classification of Landsat ETM+ imagery (81%). However, the confusion shown in some of the low user's (e.g. 20 %) and producer's (e.g. 44 %) substantial improvement yet to be made suggested by statistic.[11]

12. WORLD-WIDE GEOTAGGED IMAGEDATASET IS COLLECTED FOR AUTOMATIC IMAGE ANNOTATION AND REVERSE GEOTAGGING

In this paper, a dataset of geo-tagged photos on a world-wide scale is collected through the website called Flickr. The dataset contains a sample of more than 14 million geo-tagged photos obtained from Flickr with their corresponding metadata. The number of users upload their photos, using keywords called tags or captions and share them with each other has increased. Photos are assigned location information, i.e., geo-tagged. A geo-tagged images consists of the longitude and latitude of the location of image .Geo-tags can be automatically added to the EXIF Nowadays modern camera and phones have the feature of GPS It is also possible to assign location information manually, In this project the crawl geo-tagged photos is done on the basis of keyword search (e.g. city names).a crawling strategy aims in gathering photos from Flickr using Flicker API, so that the spatial distribution of the data is preserved. That means, the photos which are collected from a given place should reflect the popularity of that place. The crawling method starts by randomly selecting a photo from the pool of Flickr photo. The uploader of that photo is identified and the corresponding geo-tagged photos are downloaded with the corresponding metadata. To crawl more data, the complete process is repeated by selecting a new photo identifier. Furthermore, user-provided tags can be made noise-free with the help of automatic tag cleaning approach. For efficient retrieval, photos in the dataset are indexed based on their location with the help of quad-tree data structure. The dataset can be used in assisting different applications, especially, reverse geo-tagging and search-based automatic image annotation.[13]

13. UPDATED MAP OF KOPPEN-GEIGER CLIMATE CLASSIFICATION

Koppen-Geiger climate classification is the most frequently used climate classification which was developed by the German Russian climatologist Wladimir Köppen in 1884; Köppen was a plant physiologist and understood that plants are indicators for climatic elements. This effective classification was constructed on the basis of five vegetation groups (A), the arid zone (B), the warm temperate zone (C), the snow zone (D) and the polar zone (E) In 1961 it was been modified by Rudolf Geiger. Two global data sets of climate observations were used to update the historical world map of the Köppen-Geiger climate classes. Both can be seen on a regular 0.5 degree latitude/longitude grid. The first data set is provided by the Climatic Research Unit (CRU) of the University of East Anglia and the second data set (BECK et al., 2005) is provided by the Global Precipitation Climatology Centre (GPCC).[14]

14. SCIKIT-LEARN:MACHINE LEARNING IN PYTHON

Python programming language is one of the most popular languages for scientific computing. Scikit-learn is a Python module which is integrated through a wide range of state-of-the-art machine learning algorithms. Scikit-learn differs from other machine learning : i) it is distributed under the BSD license ii) it incorporates compiled code for efficiency iii) it depends only on numpy and scipy to facilitate easy distribution, and iv) it focuses on imperative programming, unlike pybrain which uses a data-flow framework while the package is mostly written in Python, it inherit the C++ libraries LibSVM and LibLinear that provide reference for the implementations of SVMs and generalized linear models with compatible licenses. Scikit-learn come across variety of machine learning algorithms including both supervised and unsupervised. Importantly, the algorithms which are implemented in a high-level language can be used as building blocks for approaches specific to a use case Future work includes online learning, to scale to large set of data.[15]

15. FOR THE IDENTIFICATION OF LAND COVERAGE IN ECOSYSTEM DONE USING IMAGE SEGMENTATION AND DISCRIMINANT ANALYSIS

The representation of the textured nature of most natural land cover units can be done in remotely sensed images. This results in the classification of per-pixel. The segmentation algorithm, Iterative Mutually Optimum Region Merging (IMORM), is presented and is used in image partition thereafter it is classified by Linear Canonical Discriminant Analysis. The per-segment approach can provide much higher accuracy .Compared to that, conventional per-pixel approach provides lower accuracy.[12]

16. CLASSIFICATION OF RANDOM FOREST ON MULTISOURCE GEOGRAPHIC DATA :

In this paper, random forest is used for classifying multisource data. Random forest is the classifier that grows many classification trees. A bootstrapped sample is used to train each tree of the training data, and at each node the algorithm only searches across a random subset of the variable to determine a split. Ensemble classification methods train various classifiers and use a voting process to combine their results. The most widely used ensemble methods are bagging and boosting. Bagging is used to reduce the variance of a decision tree classifier. Here the objective is to create several subsets of data form training sample chosen randomly with replacement. Each collection of subset data is used in the training of their decision. Boosting is used in the creation of a collection of predictors. In this technique, learners are learned sequentially with early learners fitting simple models to the data and then analysis of data is done for errors. Boosting generally reduces both the variance and the bias of the classification and has been showed to be the most accurate method. However it has many drawbacks, it is very slow, it over-train, and is sensitive to noise. Random forest is compared to boosting; they are computationally less than boosting. In experiments, the random forest classifier and was comparable to accuracies.[16]

17. CLASSIFICATIONS BASED ON HIGH RESOLUTION MULTI-SPECTRAL SATELLITE DATA FOR IMPROVING LAND

The region of rapid change in the reflectance of vegetation in the near infrared range of the electromagnetic spectrum can be called Red Edge. The red edge channel is used in the improvement of the classification of land, as the electromagnetic spectrum is sensitive for vegetation chlorophyll content. Rapid eye is the first satellite which operationally provides a red edge channel. The objective is to test the potential of the RapidEye red edge channel for improvement of land classification. Results increase the accuracy. Highest positive effects are observed for vegetation classes located in open landscapes, e.g. for bush vegetation. The data used consists of RapidEye satellite images as well as digital biotope maps and selective vegetation field mapping for ground evaluation.[17]

18. CLASSIFICATION ON URBAN VEGETATION: USING RAPIDEYE SATELLITE DATA

Due to global climate change led to an increase in number of urban dwellers and often being augmented by an aging and more sensitive population. Urban studies, still lack in classifying the urban vegetation and adequate details and across large areas. To remedy this gap, a support vector machine is used to and eight frequent tree generate is classified in the city of Berlin. Different spectral and temporal band combination of RapidEye images was investigated. The scenario clearly indicates a good classification result for tree genera using multitemporal RapidEye imagery.[18]

19. GEOLOCATION ON THE PHOTOS WITH CONVOLUTIONAL NEURAL NETWORKS

Photo geolocation is an extremely difficult task since a lot of photos provide only a few, possibly ambiguous, cues about their location. For example, the image of a beach could be taken on many coasts across the world. Even, when the landmarks are present there can still be ambiguity. Traditional computer vision programmes lack the kind of world’s knowledge, relying on the features provided by them during training. The task of geolocation can be treated as a classification problem and the subdivision of the surface of the earth into a set of geographical cells which make up the target classes can be done. We train a Convolutional Neural Network (CNN) using millions of geo-tagged photos. The resulting model is called as Planet is capable of localizing a large number of photos. Moreover, Planet is combined with LSTM (long short term memory) architecture which helps the model to achieve 50 % accuracy over the single image module.[19]

20. LEARNING DEEP FEATURES FOR DISCRIMINATIVE LOCALIZATION

It has been shown that, the convolutional units of various layers of convolutional neural network actually behave as object detectors despite no supervision on the location of the object were provided. However, this ability is lost when fully-connected layers are used for classification. Network In

Network and GoogleNet are the two popular, fully-connected neural networks that have been proposed to avoid use of fully-connected layers to minimize the number of parameters while maintaining high performance. Global average pooling is used to achieve high performance which acts as a structural regularizer, preventing overfitting during training. While the experiments going on, the global average pooling acts more than a structural regularizer. In fact, with a little tweaking effect, the network can retain its localization ability until the final layer. In a single pass, it allows to discriminate image regions for a wide variety of task. Furthermore, the localizability of the deep features in the approach can be easily transferred to other recognition datasets for generic classification, localization, and concept discovery.[20]

III. DATASET AND FEATURES

To acquire coaching and testing knowledge we tend to start by characterizing a dataset of publicly available Flickr pictures out there that were geo-tagged. We tend to then engineer a script to filter through the pictures during this dataset, supported user-provided tags. Especially, we tend to target chiefly on nature and landscape pictures for this project and that we tried to eliminate photos of individuals or urban shots. Flickr provides users with the choice to tag pictures with keywords. This allowed us to pick out pictures with keywords relevant to landscapes, and separate undesirable pictures, like portraits and street photography. Desired tags enclosed "landscape," "outdoors", "City", "People", "Mountains", "Beaches", "scenery" and "nature," whereas unwanted tags enclosed, "urban," "me," "portrait" and "bye," among others. We tend to then download the set of pictures with desired filters. This gave us a complete of concerning 3000 images, largely of natural landscapes, to figure with. Thanks to the tags being user-provided, however, some pictures didn't depict landscapes, leading to howling labels. With all, mapping the geolocations of our dataset show that we've got representative pictures of climates from most of the planet.


Figure. 1


Figure. 2


Figure. 3


Figure. 4

For the category labels we had a tendency to start using the Koppen climate organization however found that a lot of the climates were severely overshadowed by different climates with more representative pictures in our dataset. Intrinsically, we had a tendency to sort sure Koppen climates along and came up with our own broader organization for the climates. The 13 overarching categories and therefore the Koppen climates they represent are listed below in the table. Exploiting all 320,000 images, however, terminated up being too computationally hard-to-please for our resources. Thanks to this, we have a tendency to determine to solely use a sample of our initial dataset. Sure climates, like oceanic, had more pictures attributed to them, leading to a relative category imbalance, therefore in down-sampling, we selected to sample a similar variety of pictures from every climate, that balanced the categories. We felt this was cheap, as though the results aren't representative of our model's performance on the initial dataset, the model isn't meant to be used on the initial dataset, however rather user inputted pictures.

Table 1. Climate Superclass

	Superclass	Koppen Symbols
0.	Arctic/alpine	EF, ET
1.	Arid – cold	BWk, BSk
2.	Arid – hot	BWh, BSh
3.	Continental – hot	Dsa, Dwa, Dfa
4.	Humid subtropical	Cwa, Cfa
5.	Mediterranean	Csa, Csb, Csc
6.	Ocean	Ocean
7.	Oceanic	Cwb, Cwc, Cfb, Cfc
8.	Subarctic (continental - cold)	Dfc, Dfd, Dsc, Dsd, Dwc, Dwd
9.	Tropical monsoon	Am
10.	Tropical rainforest	Af
11.	Tropical savanna	Aw, As
12.	Continental – warm	Dsb, Dwb, Dfb

Figure. 5

The images were then labeled with their latitude and meridian data and Koppen & Geiger climate classification map knowledge, by finding the closest geolocation to the image data within the climate map. The element values were then traditionalized to be normal. Associate degree analysis of our final dataset showed an inexpensive unfolds of locations across the globe, in addition as across totally different climates. Because we manually balanced categories, we have to conjointly see even numbers of climates in our dataset. Finally, we have to split our dataset and stratified to take care of category balances, setting 60% for training, 20% for validation, and 20% for testing.

We have simplified the Koppen Labels dataset to retrieve the proper symbol for the climate classification for accurate result. The representation of the dataset is as follows:

Table 1. Climate Superclass

	Climate	Labels
0.	Arctic/alpine	EF
1.	Arctic/alpine	ET
2.	Arid – cold	BWk
3.	Arid - cold	BSk
4.	Arid - hot	BWh
5.	Arid - hot	BSh
6.	Continental - hot	Dsa
7.	Continental - hot	Dwa,
8.	Continental - hot	Dfa
9.	Humid subtropical	Cwa
10.	Humid subtropical	Cfa
11.	Mediterranean	Csa
12.	Mediterranean	Csb
13.	Mediterranean	Csc

14.	Ocean	Ocean
15.	Oceanic	Cwb
17.	Oceanic	Cwc
18.	Oceanic	Cfb
19.	Oceanic	Cfc
20.	Subarctic (continental - cold)	Dfc
21.	Subarctic (continental - cold)	Dfd
22.	Subarctic (continental - cold)	Dsc
23.	Subarctic (continental - cold)	Dsd
24.	Subarctic (continental - cold)	Dwc
25.	Subarctic (continental - cold)	Dwd
26.	Tropical monsoon	Am
27.	Tropical rainforest	Af
28.	Tropical savanna	Aw
29.	Tropical savanna	As
30	Continental - warm	Dsb
31	Continental - warm	Dwb
32	Continental - warm	Dfb

Figure. 6

IV. METHODOLOGY

A) SIMPLE CLASSIFICATION:

To obtain the better result by extracting the geological data from the images we have downloaded geo-tagged images from the Flickr, Google and yml to create our own geo-tagged dataset of about 3000 photos from around the world which was tagged with the type, landscape, name, city, etc. which was then filtered into recommended type for landscape images only with only geolocation metadata. Basically, our primary aim from collecting the data to train our model. In first model we are expecting result of identifying the climate region of the user given image containing the geolocation or geo-tagged photo using the Koppen Classification model and Dataset created with the different characteristics using coordinates and finally map the coordinates of given image on the map which will then classified from the dataset to extract the particular symbol from the Koppen & Geiger Dataset and then match the symbol with the classes and climate in the climate label dataset which gives the exact climate region of the picture given.

The program flow of this method is as follows:

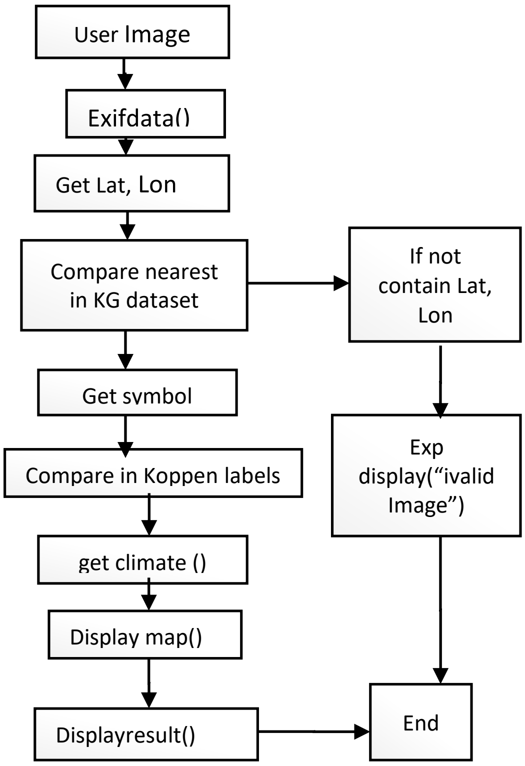

Figure. 7

B) ADVANCE CLASSIFICATION:

In order to determine a baseline for performance of a model on our dataset, we tend to train a logistical regression model and a support vector machine model. Once AN initial arrange to train these models on normalized constituent values, we tend to set to manually produce options from the pictures initial. This might cut back the spatiality of the options, and make additional tractable info. To do this, we tend to used Histograms of homeward-bound Gradients (HOG). HOG was used each for its procedure potency and its use in image classification within the literature, though not tested on landscape mental imagery specifically. HOG works by shrewd many totally different orders of gradients over the image, so shrewd frequencies of those gradients in grids across the image. Once standardization parameters, the most effective performing arts version of this formula leading to a one,568-dimensional feature vector, that was then normalized to possess mean zero and variance of one. With these options, we tend to then train a logistical regression classifier. Logistical regression models the link between options and therefore the response variable, that during this case is that the climate, through the logistical perform, that takes the form:

$$h_{\theta}(X) = \frac{1}{1 + e^{-\theta^{T} X}}$$

X is the features corresponding to a given image, and θ is the parameter that our model learns during training which we then get our prediction from by taking the SoftMax of this vector. To optimize this parameter from this output’s vectors with values of each climate in region, we perform l_2 regularization by minimizing the following cost function:

$$\min_{\theta,c} \frac{1}{2}\theta^T\theta + C\sum_{i=1}^{n} \log\left(\exp\left(-y_i(X_i^T\theta + c)\right) + 1\right)$$

where C is that the regularization term. We found, however, that once coaching the model on our derived options, we tend to achieved 100% accuracy on the coaching knowledge. Upon review, we tend to accomplished that there have been additional dimensions in our feature vector than there have been coaching examples, and therefore the model matrix wasn't full rank. To account for this, we tend to used solely the primary 100 principal elements of every feature, reworking every feature mistreatment principal element analysis (PCA). PCA finds orthogonal elements that describe the foremost variation at intervals the info, that every vector will then be projected onto. we tend to then trained AN SVM model to urge another baseline accuracy. SVM works by maximizing the price performs the function:

$$W(\propto) = \sum_{i=1}^{n} \propto_i - \frac{1}{2}\sum_{i,j=1}^{n} y^{(i)} y^{(j)} \propto_i \propto_j \left(x^{(i)}, x^{(j)}\right)$$

subject to the constraint that

$$\sum_{i=1}^{n} \propto_i y^{(i)} = 0$$

With

$$\propto_i \geq 0, \quad i = 1, \ldots., n$$

where < *x(i), x(j)* > is given by some kernel perform K. By doing this, the algorithmic program finds a hyperplane that optimally separates 2 categories of information, by increasing the minimum distance between knowledge points and also the plane. However, as a result of we've got quite 2 categories of information, we have a tendency to use a 1 vs one paradigm, making many various models to account for every pairing of categories. Finally, to tune the hyperparameters, we have a tendency to perform a grid search over regularization constants for statistical regression and kernels for SVM, victimization 3-fold cross validation on the coaching knowledge. when having our 2 baselines we have a tendency to proceeded to do and build a additional sturdy model employing a convolutional neural network. Specially, we hand-picked the ResNet-18 design. This residual neural network consists of associate degree initial convolutional layer, eight two-layer resnet blocks and a final totally connected layer. moreover, we have a tendency to use a transfer learning approach. A resnet was pre-trained on the ImageNet dataset, yielding high performance classification for general pictures.[10]

Theoretically, this permits the hidden layers of the model to already recognize helpful options concerning pictures already. We have a tendency to then take this model and modify the output and input layer for our specific classification task. We have a tendency to conjointly normalize every image to a predefined mean and variance, each slightly on top of zero, given by the initial dataset. Finally, we have a tendency to retune the model from these beginning weights on our own dataset through customary back propagation, employing a loss perform of cross-entropy. We have a tendency to found convergence of loss when coaching for thirty epochs, using 0.001 for a learning rate, a mini-batch size of 64, and momentum of 0.9 to avoid native minima. These parameters were chosen through manual standardization to reduce validation loss, because of process limitations to looking across a bigger search area.

Each residual block contains a crosscut association, by manner of adding the outputs before the block, x, to the outputs of the stacked layers, F(x), as shown in Figure a pair of. The addition of the identity is hypothesized to facilitate optimization by creating it simple for a layer to become associate degree identity mapping, by permitting F(x) to travel to zero. This effectively permits deep models to behave additional like shallower models once doing thus is additional optimum.[1]

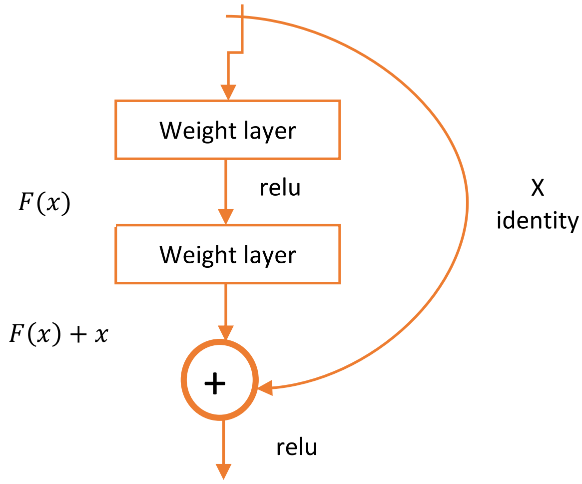

Figure. 8

V. EXPERIMENTS, RESULTS&DISCUSSIONS

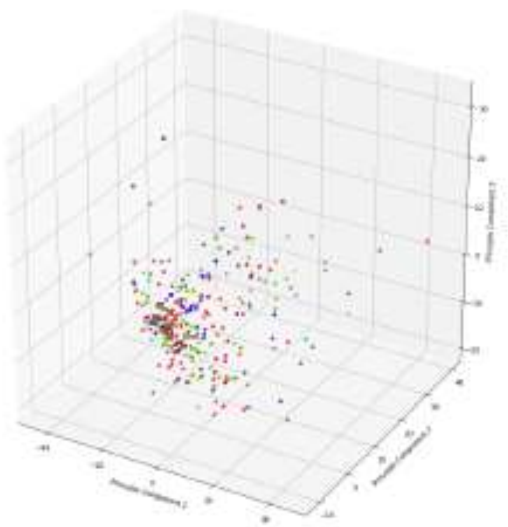

Figure. 9

To analyze our results, we tend to primarily care regarding accuracy, given just by the proper variety of classifications over the whole variety of classifications. this is often as a result of false positives and false negatives area unit adore North American nation, and wish not be weighed otherwise. Our baselines failed to perform accurately however this is often to be expected. Below may be a graph of the primary 3 parts from a PCA of the bar graph of gradients features:

As will be seen, these categories don't seem to be clearly severable during this dimension, though the primary 3 parts were solely shown to elucidate two hundredth of the variation in our options, and will still be severable in higher dimension. Clearly climate classification are a few things terribly nuanced and laborious to differentiate with the given options. Indeed, within the confusion matrix for logistical regression on the check set we tend to see poor performance.

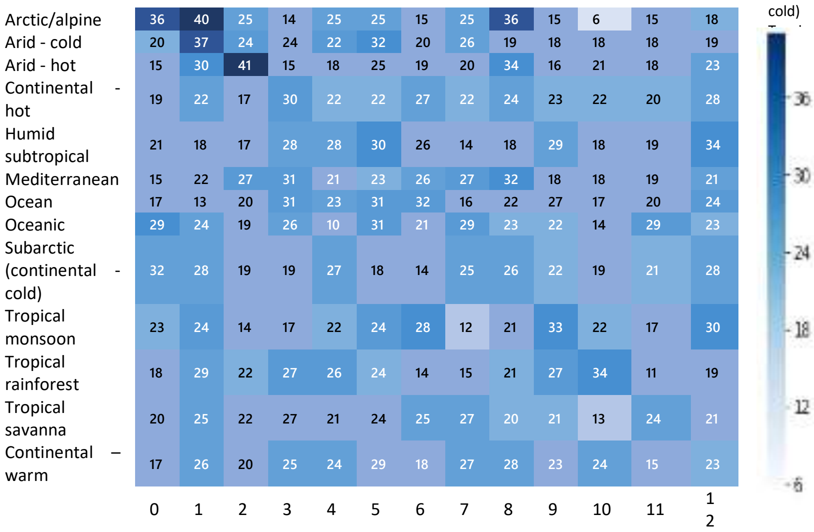

Figure. 10

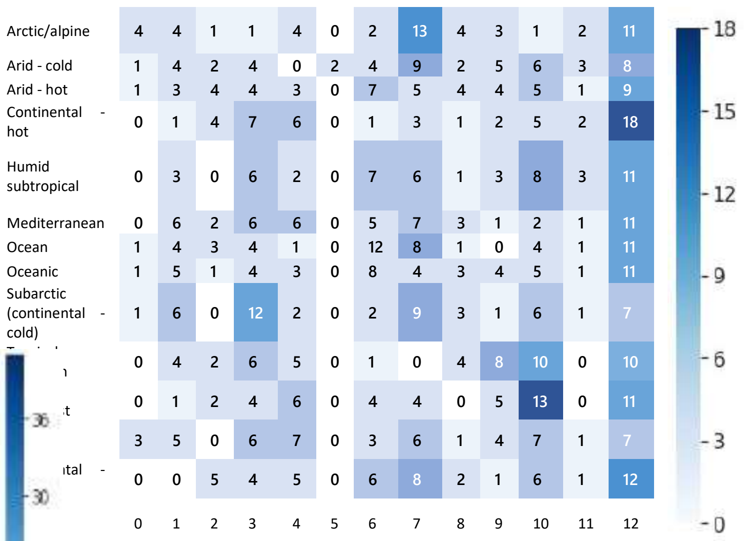

Figure. 11

Overall accuracies for the supply regression model were 0.27 on the coaching set then 0.14 and 0.2 severally on the validation and testing sets. Seeing as we've thirteen superclasses, we tend to see that the supply regression model performed solely slightly higher than probability. The SVM model conjointly performed poorly, though higher than supply regression. Whereas throughout coaching we tend to achieved associate degree accuracy of zero.76, throughout validation associate degreed testing we tend to achieved an accuracy of solely zero.13 and 0.16 severally. The confusion matrix for the performance of SVM on the testing set is reportable below.

The CNN performed higher than we expected. As is seen from the confusion matrix, even once predicting incorrectly, the anticipated categories were usually closely associated with truth category. The 3 tropical climates were usually conflated with one another, arctic and polar circle climates were joined, and also the 2 arid climates were joined still.

The table below summarizes the overall performance of each model:

	Train acc. (1610 samples)	Val. acc. (2839samples)	Test acc. (2831 samples)
LR	0.27	0.13	0.10
SVM	0.77	0.12	0.15
CNN	0.82	0.32	0.31

Figure. 12

VI. CONCLUSIONS/FUTURE WORK

To begin with, we'd have liked to have a higher administration of the dataset. As a result, we relied upon initial set of Flickr pictures with user-entered labels to classify them as landscape, nature, etc. We completed up with some noise in our dataset. Whereas we did our greatest to filter these vociferous pictures, inevitably some went through and that we complete up with pictures of interiors or shots that don't seem to be relevant for climate classification. Ideally, we thought to embrace solely nature shots. We additionally had to tag these pictures with a particular label from our thirteen super categories victimization to their GPS coordinates. Ideally, we turned up with a far better proxy for climate once labelling our dataset which might once more serve the aim of resulting in less noise or inaccuracies inside the info itself. Moreover, whereas presently our model solely takes under consideration the raw picture element values of the image itself. We have a tendency to foresee that enhancements can be created by additionally victimizing the season throughout that the photograph was taken as a feature. Landscapes will look terribly completely different looking on the time of year and this is often one thing that we saw, is inflicting misclassifications with our model in its current state. In general, snow on landscapes square measure being classified as arctic or arid circle.

VII. REFERENCES

[1] Nitze, U. Schulthess, and H. Asche, "Comparison of machine learning algorithms random forest, artificial neural network and support vector machine to maximum likelihood for supervised crop type classification," presented at Proc. of the 4th GEOBIA, 7-9 May 2012, Rio de Janeiro, Brazil (2012).

[2] P. Krahwinkler and J. Rossman, "Using decision tree based multiclass support vector machines for forest mapping," presented at IEEE Int. Geoscience and Remote Sensing Symp., Vancouver, Canada (2011).

[3] S. Adelabu, O. Mutanga, E. E. Adam, M. A. Cho, "Exploiting machine learning algorithms for tree species classification in a semiarid woodland using RapidEye image," J. Appl. Rem. Sens. 7(1) 073480, Nov. 2013. https://www.spiedigitallibrary.org/journalArticle/Download?fullDOI=10.1117%2F1.JRS.7.073 480

[4] B. Bechtel and C. Daneke, "Classification of Local Climate Zones Based on Multiple Earth Observation Data," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing , vol. 5, no. 4, pp. 1191-1202, Aug. 2012, https://doi.org/10.1109/JSTARS.2012.2189873

[5] J. C. W. Chan and D. Paelinckx, "Evaluation of random forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery," Remote Sens. Environ. 112(6), 2999–3011 (2008), http://dx.doi .org/10.1016/j.rse.2008.02.011.

[6] Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A. , Hess, K. T., Gibson, J. and Lawler, J. J., "Random Forests for Classification in Ecology," Ecology , vol. 88, pp. 2783-2792, 2007. doi:10.1890/07-0539.1

[7] D. C. Duro, S. E. Franklin, M. G. Dubé, "A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery," Remote Sensing of Environment , vol. 118, pp. 259-272, ISSN 0034-4257, 2012, https://doi.org/10.1016/j.rse.2011.11.020

[8] G. M. Foody and A. Mathur, "The use of small training sets containing mixed pixels for accurate hardimageclassification:trainingonmixedspectralresponsesforclassificationby a SVM," Remote Sens. Environ. 103(2), 179–189 (2006), http://dx.doi.org/10.1016/j.rse .2006.04.001.

[9] He, K., Zhang, X., Ren, S., & Sun, J., "Deep residual learning for image recognition," Proceedings of the IEEE conference on computer vision and pattern recognition , pp. 770-778, 2016.

[10] Navneet Dalal, Bill Triggs. Histograms of Oriented Gradients for Human Detection. International Conference on Computer Vision & Pattern Recognition (CVPR '05) , Jun 2005, San Diego, United States. pp.886--893.

[11] L. J. Quackenbush and Y. Ke, "Investigating new advances in forest species classification," presented at Proc. of 2007 ASPRS Annual Conf., 7-11 May 2007, Reno, Navada (2007).

[12] A. Lobo,"Image segmentation and discriminant analysis for the identification of land cover unitsinecology,"IEEETrans.Geosci.RemoteSens.35(5),1136–1145(1997),http://dx.doi .org/10.1109/36.628781.

[13] H. Mousselly-Sergieh et al., "World-wide scale geotagged image dataset for automatic image annotation and reverse geotagging," Proceedings of the 5th ACM Multimedia Systems Conference , pp. 47-52, Mar. 2014. DOI: 10.1145/2557642.2563673.

[14] M. Kottek et al., "World Map of the Köppen-Geiger climate classification updated," Meterol. Z. , vol. 15, pp. 259-263, 2006. DOI: 10.1127/0941-2948/2006/0130.

[15] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay. "Scikit-learn: Machine Learning in Python", Journal of Machine Learning Research , 12, pp. 2825-2830, 2011.

[16] J. A. Benediktsson and J. R. Sveinsson, "Random forest classification of multisource remote sensing and geographic data," in 2004 IEEE Int. Geoscience and Remote Sensing Symp., Anchorage, Alaska, pp. 1049–1052 (2004).

[17] C. Schuster, M. Forster, and B. Kleinschmit, "Testing the red edge channel for improving land-useclassificationsbasedonhigh-resolutionmulti-spectralsatellitedata,"Int.J.Remote Sens. 33(17), 5583–5599 (2012), http://dx.doi.org/10.1080/01431161.2012.666812.

[18] J. Tigges, T. Lakes, and P. Hostert, "Urban vegetation classification: benefits of multitemporal RapidEye satellite data," Remote Sens. Environ. 136, 66–75 (2013), http://dx.doi.org/ 10.1016/j.rse.2013.05.001.

[19] T. Weyland, I. Kostrikov, and J. Philbin, "PlaNet - Photo Geolocation with Convolutional Neural Networks," arXiv:1602.05314 [cs.CV], Feb. 2016.

[20] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A., "Learning deep features for discriminative localization," Proceedings of the IEEE conference on computer vision and pattern recognition , pp. 2921-2929, 2016.