Classifying Remote Sensing Data

Vibhore Singh
u3248455@uni.canberra.edu.au

Artificial Intelligence Techniques
University of Canberra

Abstract

Identification of the ground cover using Remote sensing holds significant potential for advancing research in biodiversity conservation, environmental monitoring and informed decision-making regarding climate change on a global scale. A substantial amount of satellite image data is being generated on a daily basis. There is a significant need for the methods that automatically classify that data based on their features. In this project, we aim to develop a classification framework for the data received from Landsat 5 satellite for the mountainous region in Colorado. . A quantitative approach was employed to find the patterns and relationships between the 6 spectral bands of our 15,000 instances to predict the corresponding target values representing 14 different land covers. Several machine learning models such as random forests,SVM and Decision Trees and deep learning techniques like CNN and ResNet were used for the task. Our analysis revealed that the three deep learning techniques show promising results with 94.3%, and 91% accuracies . This study has the potential to be crucial for studying environmental changes and supporting research in various domains, including global ground cover change, agriculture, geology, forestry, regional planning, surveillance, education etc. The evolving dataset and the proposed classification framework can assist in the creation of more robust models capable of accurate and continuous monitoring of the satellinte images. This study can provide valuable insights for making informed decisions about ecosystems and climate change worldwide.

Keywords : **Remote Sensing, Satellite Images, Image Classification, Remote Sensing Data Classification, Spectral Bands Classification, Ground Cover Classification, Machine Learning, Deep Learning, Convolutional Neural Network**

1 Introduction

The Landsat program, jointly developed by NASA and USGS in 1972, is the longest- running program for satellite imagery of Earth. It is a pioneer in acquiring satellite imagery of Earth's cover and acquires millions of images of different parts of the Earth. [1] [2] With the ever-growing volume of the satellite image data, the need for harnessing this data and use it to accurately map, delineate and characterize land cover is also increasing. However traditional ground cover classification methods are often labor intensive, time consuming, expensive and often limited in their analytical prowess. Thus the need of the hour is to use the state of the art machine learning models to automate the classification task.

Despite its huge benefits and potential, the automatic classification of the satellite images is a challenging task especially when we are dealing with the rugged terrains such as mountainous regions. This study aims to develop an innovative classification model capable of classifying the ground covers by studying the patterns among the spectral bands deduced from the images generated by Landsat-5 remote sensing program focusing on the mountainous region of Colorado. Leveraging a quantitative research approach and the advancements in data science and machine learning, we aim to analyze the spectral information captured by the satellite's imagery to accurately classify ground cover data.

LANDSAT Colorado data set is an image of the mountainous region in central Colorado with the pixel size 1907×1784 1 and segments are extracted from the image. The sensors of LANDSAT produce images of the ground with seven different spectral bands. The resolution of the band $1-5$ and 7 is $30 * 30\ m^2$ and of band 6 it is $60 * 60\ m^2$. Thereby, the first three bands are in the visible spectrum (corresponding roughly to the colors blue, green and red) and the bands 4, 5, 7 are in the near-infrared spectrum. Thermal band 6 has a lower spatial resolution and therefore, it was dropped it according to common practice [3]. The bands are designed to detect and distinguish between different vegetation, rock formations, and other ground characters [9]. The dataset contains 3.402 million data samples with only 15,000 labelled.

The labelled dataset is provided Dr. M. Augustijn (University of Colorado) [3]. In this project, we formulate a robust classification framework on the spectral bands and classify every instance among the 14 ground cover labels. Table 1 shows the class distribution of the different classes in our complete dataset(3.4million data samples). We employ a diverse set of algorithms including random forests, support vector machines (SVM), decision trees,convolutional neural networks (CNN), recurrent neural networks (RNN), and residual networks (ResNet). We aim to evaluate the performance of these models to identify the most effective approaches for accurately classifying ground cover in mountainous landscapes.

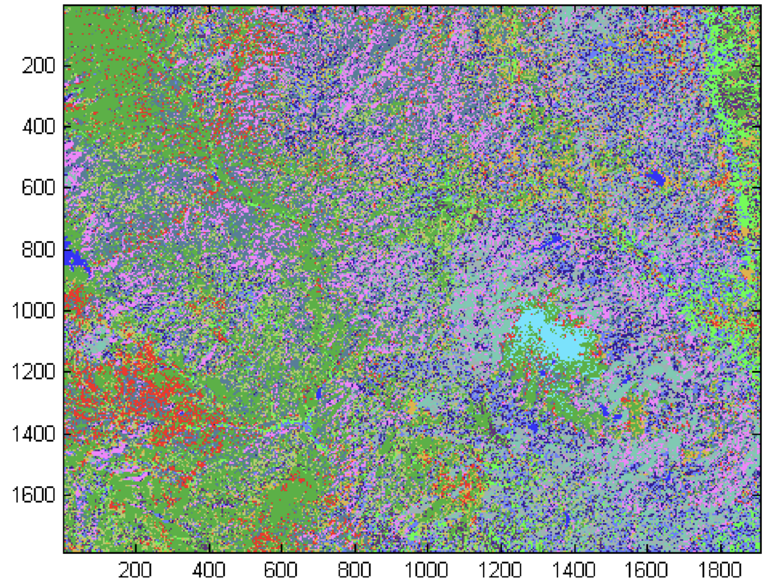


Figure 1: Colorado data set: False color image of a mountainous region in Colorado

In this paper, we make the following contributions:

1. We try to uncover the patterns between spectral bands and the ground cover types which can help in future research of the ecosystem.

2. We propose a robust classification framework for identifying the ground cover in the

mountainous regions

3. We demonstrate the combination of machine learning techniques with remote sensing data can help in increased accuracy and efficiency

4. We plan to find the most optimal technique for the classification task on the basis of accuracy.

The project represent a significant step in advancing remote sensing data classification, ultimately contributing to the understanding and management of natural resources, ecosystems, and the impacts of climate change on a global scale. This study can assist environmental agencies and policymakers by providing accurate land mapping based on satellite images for better resource management and informed policy development.

Class	Label	Distribution (%)
1	Scotch pine	17.1
2	Douglas fir	10.4
3	Pine/fir	5.3
4	Mixed pine fir	8.0
5	Supple/prickle pine	4.3
6	Aspen/mixed pine forest	6.1
7	Without vegetation	5.0
8	Aspen	8.2
9	Water	0.5
10	Moist meadow	2.9
11	Bushland	3.7
12	Grass/pastureland	7.8
13	Dry meadow	9.8
14	Alpine vegetation	10.8

Table 1: Class distribution of different ground cover classes.

2 Literature Review

With the advancements in Machine Learning and Computational technologies, Satellite image classification underwent a significant change in the last few decades. We are clearly moving from manual interpretation and analysis towards a more automated methods due to the ever increasing amount of data generated on a daily basis.

The creators of this dataset Dr. Augusteijn *et al*.(1993) in [3] investigate various feature selection methods for classifying ground covers in LANDSAT images using neural networks. They compared the traditional statistical methods and complex features with neural based techniques. They found out that combining feature from multiple spectral bands resulted in comparable performance of the models showing the capability of neural networks to handle multi spectral data. Their findings highlighted that neural networks are well adept at managing spectral data and demonstrated the potential of neural networks to simplify complex multi-dimensional data.

Helber *et al*.(2019) [7] introduced the Eurosat dataset based on Sentinel-2 satellite data spread across the 13 spectral bands. They used deep CNN for the classification and achieved the accuracy of 98.57%. They not only showed us the efficiency of deep learning models in the remote sensing data classification but also set a benchmark for the future research.

Castelluccio *et al*. [5] explored the use of deep learning further in the area of land use and land cover classification using a Sentinel-2 dataset and mainly focusing on the role of different spectral bands. They used CNN architecture like RES NET and GoogleNET and performed fine tuning. They showed that the CNN pre trained model receive high accuracy (over 90%) and thus outperforming other machine learning models. They showed the advantages of using CNN and their ability to learn without the need of extensive feature engineering traditionally required by most of the models.

In their study, Gong Cheng *et al*. in [6] , focused on enhancing remote sensing image scene classification by proposing a new large scale benchmark dataset. They compared the performance of traditional machine learning and deep learning datasets and demonstrate that deep learning methods outperform traditional classification techniques and further fine tuned the models to leverage full potential on the ever increasing data.

The paper [8] , introduces the Canadian cropland dataset containing 78,536 high resolution images representing ten crop classes over the span of four and a half years . Each image features multiple spectral bands and advanced deep learning techniques such as ResNet, DenseNet and 3d-CNN are used for the classification purposes. The study also shows the importance of deep learning techniques as they outperform traditional classification techniques.

In the paper [12], Wessel *et al*. tried to use various machine learning models on the freely available Sentinel 2 data containing spectral bands with different resolutions to classify the tree species across the two forests in Germany. Along with that they used seasonal data to show the classification of coniferous vs. broad-leaved trees. SVM showed the highest accuracy with 97%.

L B Yeo *et al*. in [13] discussed and integrated framework to show the land cover changes in Kelantan, Malaysia, using both supervised and unsupervised classification techniques applied to Landsat 8 satellite data. The study uses both supervised techniques and unsupervised techniques with the supervised model showing higher frequency that the ICUC clustering model.

While reviewing [7] and [12], we saw that they achieved the accuracy of 98% and 97% by using CNN and SVM respectively. These results informed us about the potential of these models , prompting the use of similar models in our classification task. [5] used CNN, ResNet for the land cover classification with high accuracy and [8] used deep learning for the agricultural dataset , explaining how the need for feature engineering is reduced when dealing with the deep learning techniques, which help us to adapt the similar methodology for our project.

Research Gap

During my intensive Literature Review I realised there hasn't been many studies done on the rugged,mountainous terrains presenting different challenges that those in [5], [7] and others predominately focused on more uniform landscapes and [8] and [12] majorly focused on the agriculture and tree covers. We, with our 14 different classes, try to distinguish different parts of the image like different kinds of trees(like Scotch pine , Douglas fir, Mixed pine fir etc.) , No vegetation lands, Water bodies, Meadows(moist and dry), shrubs and bushes etc. in the mountainous region of Colorado as the previous researches may not adequately represent the complexities found in the mountainous regions.

By employing a variety of machine learning and deep learning techniques, this project not only tests the effectiveness of these models in a new context but also enhances the research on land cover mapping in rugged areas and hilly terrains of the Colorado region.Our study not only bridge the research gap but also serve as a benchmark for future research in environmental monitoring, ground use planning,conservation efforts, biodiversity etc.

3 Methodology

3.1 Dataset Loading and Preprocessing

The study utilized the dataset of the spectral bands of the image taken from Landsat V satellite each sample containing values of spectral bands.Image segments were extracted from the image in different categories. [3]. This study used six of the seven frequency bands that are standard. The data samples were associated with the specific type of vegetation, ground bodies and land cover. The labelled data set has with 14 very unbalanced classes as seen in figure **??**.

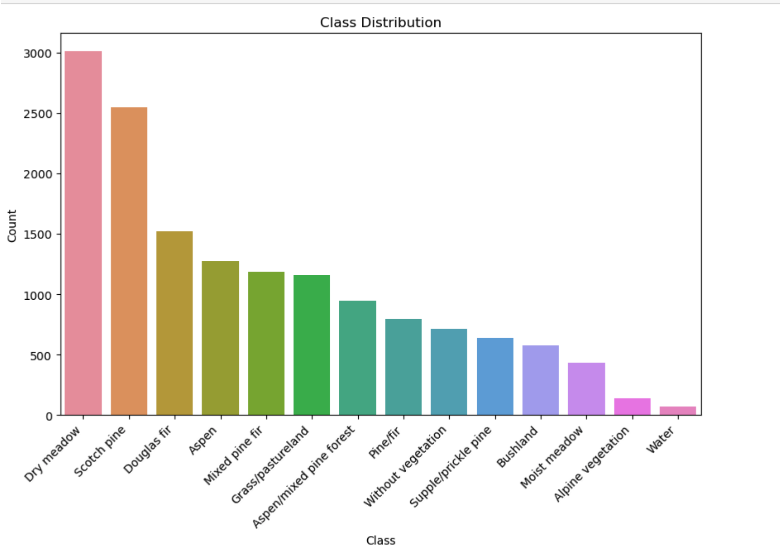


Figure 2: Distribution of Classes in the labelled data

3.2 Data Analysis

Data was loaded into a Python environment using the pandas library. Given the nature of the raw dataset, which was without headers, columns were manually labeled as $Band1, Band2, Band3$ for the spectral bands, along with *Class* for the target variable. Then the checks for missing values and Outliers was performed for every spectral bands. Then we split the data using 80:20 ratio.

Next the correlation between the dependent and independent variables. To standardize the input features and ensure that they contribute equally to the analysis Normalisation was done on the bands data. For deep learning part we reshaped the data.

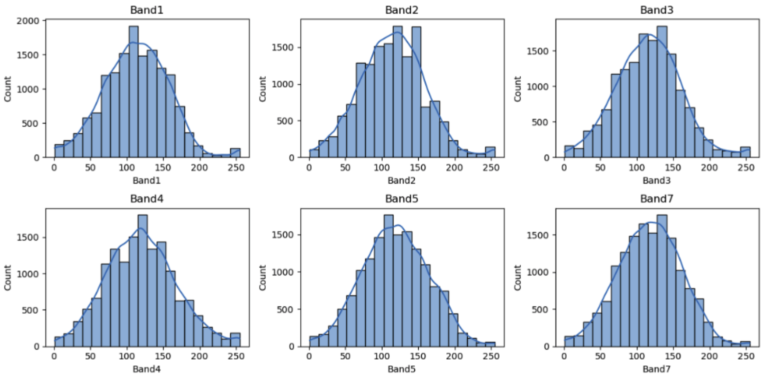


Figure 3: Distribution of the 6 bands

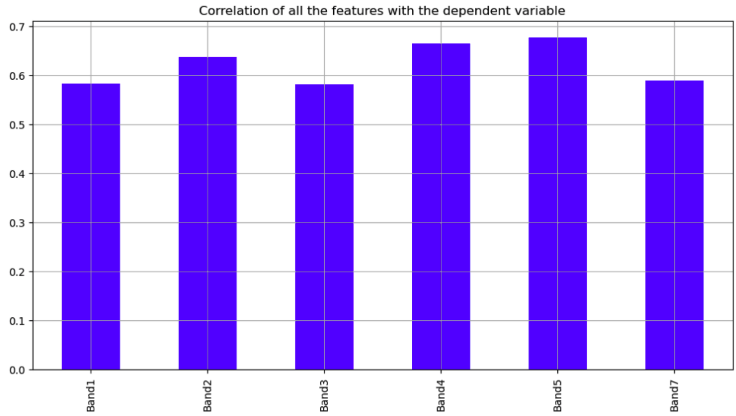


Figure 4: Correlation of all the features with the dependent variable

3.3 Model Development

This sction elaborates on the model development phase of our research which uses three traditional machine learnng models and three advnced deep learning models we thought fit for Multispectral image classification. The six model were carefully chosen to exploit different aspects of the data, from its spatial composition to its sequential patterns.

The Decision Treeand Randon Forests are the tree based models known for their robustness against noise and ability to handle unbalanced datasets, which was one of the major concerns with our dataset.RF is also known to reduce overfitting among its trees. SVM has proven effective in several high dimensional classification tasks like image classification. It helps define decision boundary with fewer errors.

CNN are ideal for the image data analysis due to their capability to learn the patterns and hierarchies among the spatial data through their convolution filters. This makes them one of the best choices and thus so many researchers have used it for the classification of remote sensing images.[8] [7]. Long Short-Term Memory (LSTM) Networks are used to handle the sequential data problems and it seems fit for our dataw which has the seqential arrangement of the spectral bands. Resnet is a pretrained model with the ability to train very deep networks using residual blocks with skip connections and jumping over layers which helps them to learn more complex patterns and features at various scales, which is crucial for accurate classification.

We initiated with the default parameters of the traditional ML models to get an understanding of how our models are working and then we prformed hyperparameter tuning to find the best perfoming model with the best parameters to get the best classification results. For the deep learning models we also initiated with the basic models but made changes to get better results . We used hit and trial method to find the optimal number of layers for each model.

CNN Architecture

The CNN model we finalised has two convolutional layers with 32 and 64 filters respectively, a kernel size of 3, 'relu' activation, This layer helps in capturing the spatial hierarchy of features. The convolution layers are followed by Pooling layers to reduce the spatial dimensions. Next comes a Flatten layer which takes the output from the convolution and pooling layers to form a 1d vector. Adense layer to learn non-linear combinations of the high-level features extracted by the convolutional layers. A dropout layer with 0.5 rate to prevent overfitting. Then a last output layer with a softmax activation function that classifies the input into multiple categories based on the number of target classes.

1. Convolutional Layer 1: 32 filters with a kernel size of 3 and ReLU activation.
2. Pooling Layer 1: Max pooling with a pool size of 2
3. Convolutional Layer 2: 64 filters with a kernel size of 3 and ReLU activation.
4. Pooling Layer 2: Max pooling with a pool size of 2.
5. Flatten Layer: Flattens the output of the previous layers into a 1D array.
6. Dense Layer: A fully connected layer with 128 neurons and ReLU activation.
7. Output Layer: A softmax layer with a number of neurons equal to the number of classes, providing probabilities for each class.

LSTM
The LSTM model has 2 LSTM layers The first layer returns sequences meaning the output is preserved for the next layer and a second layer whichdoes not return sequence. Next are the fully connected dense layer with 100 units and 'relu' activation and a dropout layer with a rate of 0.5 to reduce overfitting. The Output Layer uses a softmax activation function for classification across the target variable.

ResNet
For Resnet we used very basic model which utilized residual blocks extensively to enable the training of deeper networks without degradation, helps in recognizing intricate patterns and relationships between the spectral bands.

3.4 Model Selection

Next step was to find the best performing model for the task. We used grid search cross validation with different values of parameters for each of the DT, SVM and RF. For SVM we tire with several values of c, Kernel, Gamma whereas for RF and Dt we use Max depth, Min samples split, no of trees for RF. Out of these three SVM was th best performing algorithm on the basis of accuracy and F1 score anf gave us the best parameters as 'C': 100, 'gamma': 'scale', 'kernel': 'rbf'.

For deep learning algorithms we fiddled with the layer depths , kernel size drop out rate(for CNN) and Number of Units, Batch Size, and number of layers (for LSTM). Next we implemented a learning rate scheduling strategy to enhance the training process of the models. As the learning rate decreases, the minimum of the loss function become smaller, allowing more refined adjustments in the model weights, which can lead to better overall performance and prevent over fitting as well.

Based on metrics such as Accuracy , F1 score etc we decide to choose CNN model as our final model . Even though SVM has slightly better accuracy than CNN but due to the fact that CNN are basically designed to perform image analysis and the amount od data we have, we decided to go with the best performing deep learning dataset as we know deep learning is better for larger dataset. CNN are well adept at capturing spatial patterns and are robust to the variation in the input data.

CNN

CNN are the type of deep learning algorithms used for images and other grid based data analysis. It's also known as a ConvNet. A convolutional neural network is used to detect and classify objects in an image. [10] It identifies the spatial hierarchy of data using its several components:

1. **Convolution Layer :** These layers apply a number of filters to the input. This layer performs a dot product between two matrices, where one matrix is the set of learnable parameters otherwise known as a kernel, and the other matrix is the restricted portion of the receptive field. [10]The kernel is spatially smaller than an image but is more in-depth. Each filter transforms a part of the input , defined by the kernel size, using the dot product between the filter and local regions of the input, to produce a feature map. This process allows the model to capture small, local features of the input. [11]Mathematically,

$$(I * K)(i, j) = \sum_m \sum_n I(i - m, j - n) \cdot K(m, n)$$

Here, I is the input matrix, K is the kernel, and (i, j) are the coordinates in the output feature map. m and n the dimensions of the kernel.

2. **Activation Functions:** After each convolution operation, an activation function is applied to introduce non-linearities into the model. In simple words it sets the negative pixels to 0 helping it learn more complex patterns. A common activation function used in CNNs is ReLU (Rectified Linear Unit).[4]

3. **Pooling Layers:** Pooling is a down-sampling operation that reduces the dimensionality of the feature map and reducing the computational complexity. The rectified feature map goes through a pooling layer to generate a pooled feature map. Mathematically ,[11][4]

$$P(i, j) = \max_{m,n \in \text{region}} I(i + m, j + n)$$

Here, P is the output of the pooling layer, I is the input feature map, and the max operation is taken over a specified region (typically 2x2 or 3x3).

4. **Fully Connected Layers:** After several convolutional and pooling layers, the high-level reasoning in the neural network is done via fully connected layers. Neurons in this layer have full connectivity with all neurons in the preceding and succeeding layer as seen in regular FCNN.[10]

$$y = \sigma(Wx + b)$$

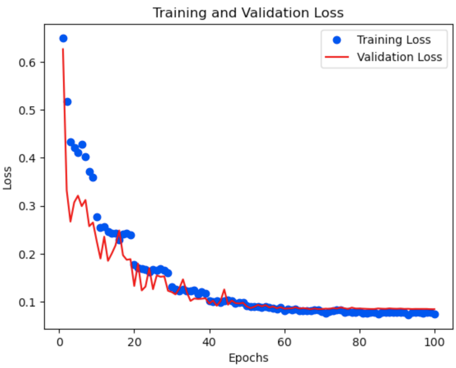
Where x is the input vector to the layer, W is the weight matrix, b is the bias vector, σ is an activation function, and y is the output vector.

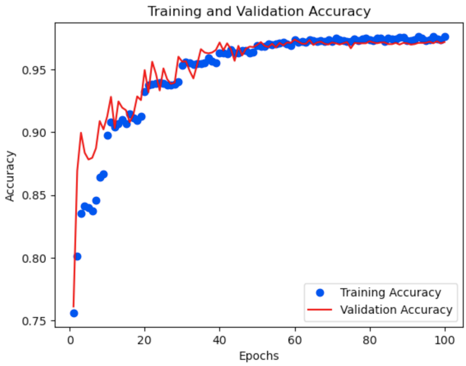
4 Evaluation:

The next step is the evaluating the performance of our models. The metrics we use for to evaluate are Accuracy, Precision and Recall and the F1 score. The table below 2 gives the values of accuracy and and f-1 score for our initial and final models . This will help us in selecting CNN as our best models.

Model	**Accuracy**		**F1 Score**	
	Initial	**Final**	**Initial**	**Final**
RF	0.93	0.93	0.92	0.92
DT	0.89	0.89	0.87	0.87
SVM	0.95	0.97	0.94	0.96
CNN	0.95	0.97	0.95	0.96
LSTM	0.93	0.96	0.91	0.94
ResNet	0.89	0.89	0.87	0.87

Table 2: Evaluation of Machine and Deep Learning Models





Visualisations :

1. **Training vs Validation Accuracy Graph** :

    - The training accuracy shows the proportion of correct predictions whereas the validation accuracy shows how well the model generalizes to the new data.
    - The graph shows the accuracy starts quite low and then rapidly increases suggesting the model is learnin on the data.
    - The model is learning effectively during the initial epochs.
    - The training and validation accuracies converge closely, which is a good indicator of the model's generalization capabilities
    - Both the accuracies stabilizes around 95% indicating good model performance
    - The graph maintains a stable performance over most epochs, which is ideal for dependable model deployment, showing that the learning process is robust.

2. **Training vs Validation Loss Graph:**

    - Loss is a measure of how well the model's predictions match the actual labels
    - Lower the loss the better the model is.
    - Both training and validation loss decrease sharply at the start, demonstrating effective learning and optimization of the model weights in response to the training data.
    - over time the loss begins to decrease as a very decreased rate showing the model is near convergence.
    - The validation loss shows slightly more variance than the training loss

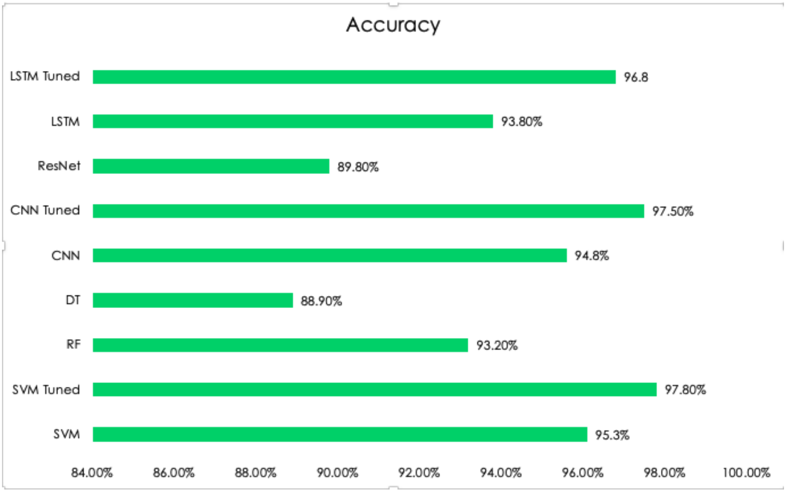


Figure 5: Accuracy comparison of different models

```
CNN Model - Classification Report:
              precision    recall  f1-score   support

           0       0.99      0.97      0.98       491
           1       0.95      0.99      0.97       298
           2       0.92      0.93      0.93       168
           3       0.98      0.92      0.95       241
           4       0.94      0.97      0.96       135
           5       0.96      0.98      0.97       196
           6       0.95      0.79      0.86       156
           7       0.92      1.00      0.96       240
           8       1.00      0.72      0.84        18
           9       0.92      0.93      0.92        85
          10       0.99      0.85      0.91       111
          11       0.85      1.00      0.92       225
          12       0.98      0.97      0.98       600
          13       0.97      0.94      0.96        36

    accuracy                           0.95      3000
   macro avg       0.95      0.93      0.94      3000
weighted avg       0.96      0.95      0.95      3000

CNN Model - Weighted F1 Score: 0.9536120251592726
```

Figure 6: CNN

```
Classification Report (SVM):
              precision    recall  f1-score   support

           1       0.98      0.99      0.98       491
           2       0.96      0.99      0.97       298
           3       0.94      0.96      0.95       168
           4       0.98      0.93      0.95       241
           5       0.94      0.93      0.93       135
           6       0.96      0.97      0.97       196
           7       0.99      0.79      0.88       156
           8       0.95      0.97      0.96       240
           9       1.00      0.78      0.88        18
          10       0.96      0.96      0.96        85
          11       0.93      0.95      0.94       111
          12       0.92      0.96      0.94       225
          13       0.97      1.00      0.98       600
          14       0.97      0.94      0.96        36

    accuracy                           0.96      3000
   macro avg       0.96      0.94      0.95      3000
weighted avg       0.96      0.96      0.96      3000

SVM F1 score: 0.9471752470839393
```

Figure 7: SVM

```
LSTM Model - Classification Report:
              precision    recall  f1-score   support

           0       0.98      0.96      0.97       491
           1       0.95      0.98      0.97       298
           2       0.95      0.95      0.95       168
           3       0.96      0.97      0.96       241
           4       0.94      1.00      0.97       135
           5       0.95      0.95      0.95       196
           6       0.99      0.85      0.91       156
           7       0.92      0.99      0.96       240
           8       0.93      0.72      0.81        18
           9       0.97      0.92      0.95        85
          10       0.95      0.92      0.94       111
          11       0.93      0.93      0.93       225
          12       0.98      0.99      0.99       600
          13       1.00      0.92      0.96        36

    accuracy                           0.96      3000
   macro avg       0.96      0.93      0.94      3000
weighted avg       0.96      0.96      0.96      3000

LSTM Model - Weighted F1 Score: 0.9606786807451463
```

Figure 8: LSTM

Figure 9: Classification reports of top three performing algorithms

4.1 New data classification

In the development of our Convolutional Neural Network (CNN) model, the objective was to achieve efficient and accurate classification of remote sensing imagery characterized by multiple spectral bands. We were able to achieve the classification on the labelled data with 97% accuracy. The model architecture is explained in Section 3.

Once we have the final CNN model , we need to apply the model on new data to find the class labels for the unlabelled dataset. In our unlabelled dataset we have 3.4 million data samples. We name the columns and used the similar scaling and reshaping techniques for this data as well. Finally the prepared data was fed to the CNN model for predictions and it gives us the predictions for the unlabelled data. Figure 10 gives the distribution of predictions made by the model

4.2 Misclassifications :

Even though classes such as Douglas fir, Scotch pine and pine/fir were classified approximately accurately as compared to Table 1 We can see there are some mis classifications. Class 12 representing Dry Meadow is classified as the most frequent class but there are other

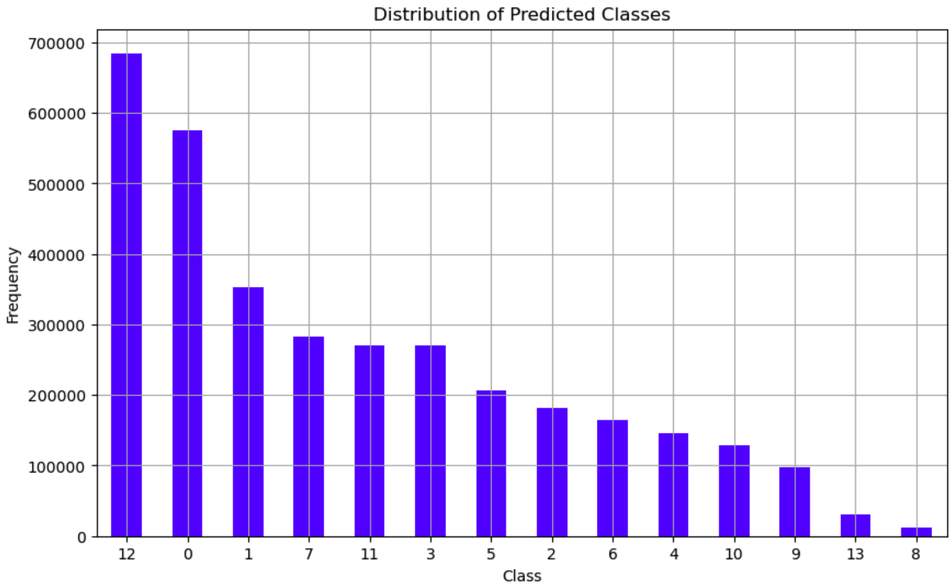


Figure 10: Distribution of classes in unlabelled data

classes with higher proportion than it. Class 13 represent Alpine vegetation which was supposed to be 10.8% but in the graph we can see it is one of the lease classified class.

The reasons behind these misclassification are:

1. The labelled data we trained our model on was not even 1% of the unlabelled data. Since the model learns on such a small data it doesn't give proper classifications.

2. The unlabelled data classifications are in line with the proportion of the classes present in labelled data.

3. The class imbalance lead to model biases, where the model performs well on majority classes but poorly on minority classes..

5 Future Work

5.0.1 Learnings and Insights

1. Multiple classes Classification in the large dataset present unique challenges primarily related to class imbalance.

2. CNNs seem highly efficient in managing multi-spectral data inputs, where each spectral band provides different information about the surface

3. We found out that CNN proved to be the most efficient and best performing model in the task of remote sensing data classifications.

4. Deep learning models tend to work better than the traditional machine learning models.

5. Deep learning models automate much of the feature detection required for high-performanc results

6. A key learning from our project was the dependency of deep learning models on large volumes of diverse training data to function optimally.

7. There are misclassifications on the new data if the size of training data is too small.

5.1 Future Research Directions

1. **Necessity for Larger Datasets:** For training deep learning models like CNNs, a larger and more diverse dataset is crucial. Increasing the learning dataset size can improve the model's ability to learn subtle patterns that are often critical for accurate classification in complex environments like remote sensing.

2. **Complex Architectures and Hybrids:** Exploration of advanced complex deep learning architectures and CNN hybrids may lead to improvements in model accuracy and efficiency by effectively capturing both spatial and temporal dependencies, or by enhancing feature extraction capabilities.

3. **Transfer Learning**Leveraging transfer learning techniques for this task can significantly decrease the computational cost and time for the deployment for new models. By using pre-trained models on similar tasks and fine-tuning them we can speed up the training process and improve model performance, especially when labeled data are scarce.

4. **Feature Engineering:** Despite the strong feature extraction capabilities of CNNs, manually selected features through traditional engineering techniques can still enhance model performance.

5. **Importance of Balanced Training Data:** Ensuring that the training dataset is balanced across different classes can help reduce the model bias and assist in developing robust CNN models

References

[1] United states geological survey. https://www.usgs.gov/centers/eros/science/usgs-eros-archive-landsat-archives-landsat-4-5-themat qt-science_center_objects=0#qt-science_center_objects, .

[2] Landsat program. https://en.wikipedia.org/wiki/Landsat_program, .

[3] Marijke F. Augusteijn, Kelly A. Shaw, and Rebecca L. Watson. A study of neural network input data for ground cover identification in satellite images. In *Proceedings of the International Conference on Artificial Neural Networks (ICANN '93)*, pages 1–8. Springer, 1993.

[4] Avijeet Biswal. Convolutional neural network tutorial. https://www.simplilearn.com/tutorials/deep-learning-tutorial/convolutional~neuralnetwork#:~:text=A%20convolutional%20neural%20network%20is,classify%20objects%20in%20an%20image.

[5] Marco Castelluccio, Giovanni Poggi, Carlo Sansone, and Luisa Verdoliva. Land use classification in remote sensing images by convolutional neural networks. In *DIETI, Universit'a Federico II di Napoli, Naples, Italy.*, 2015.

[6] GonG ChenG, Junwei han, and XiaoqianG Lu. Remote sensing image scene classification: Benchmark and state of the art. *Proceedings of the IEEE*, 96(8), 2008.

[7] Patrick Helber, Benjamin Bischke, Andreas Dengel, and Damian Borth. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *arXiv preprint arXiv:1709.00029*, 2019.

[8] Amanda A. Boatswain Jacques, Abdoulaye Baniré Diallo, and Etienne Lord. The canadian cropland dataset: A new land cover dataset for multitemporal deep learning classification in agriculture. *Journal of Agricultural and Environmental Monitoring*, 1(1): 1–25, 2023.

[9] Marika Kaden. Integration of auxiliary data knowledge in prototype based vector quantization and classification models. University of Leipzig, Faculty of Mathematics and Computer Science, Dissertation, 2016.

[10] Mayank Mishra. Convolutional neural networks, explained. https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939.

[11] Pranshu Sharma. Basic introduction to convolutional neural network in deep learning. https://www.analyticsvidhya.com/blog/2022/03/basic-introduction-to-convolutional-neural-network-in-deep-le

[12] Mathias Wessel, Melanie Brandmeier, and Dirk Tiede. Evaluation of different machine learning algorithms for scalable classification of tree types and tree species based on sentinel-2 data. *Remote Sensing*, 10(9), 2018.

[13] L B Yeo, W S N Wan Mohamad, R Hasan, N I Othmani, N H Abdul Hamid, N Ramlee, and O T S Yeo. An integrated framework for assessing land-use/land-cover of kelantan, malaysia: Supervised and unsupervised classifications. *IOP Conference Series: Earth and Environmental Science*, 1053(1):012026, 2022. doi: 10.1088/1755-1315/1053/1/012026.