Machine Learning Model For Detecting Lineas of Europa

Pradeep Kumar Baaisakh* and Anshuman Panda†

Guides: Dr.Guneshwar Thangjam, Dr.Subhankar Mishra

National Institute of Science Education and Research, HBNI, Jatni-752050, India

Europa, Jupiter's icy moon, captivates with its unique geology, hosting a vast ocean beneath its icy exterior, sustained by Jupiter's tidal forces. The red scars, Lineas, add intrigue, their crimson hue sparking questions about Europa's geophysical processes. Discovering this ocean challenges traditional habitable zone concepts, expanding the search for life beyond star-centric boundaries. The project focuses on Lineas, aiming to unravel their mysteries, examine salts, and potentially uncover organic compounds. Recent detection of hydrated salts heightens anticipation, promising valuable insights into Europa's composition. This enigmatic moon holds the potential to reveal cosmic secrets, making it a compelling celestial target.However detecting lineas is not an easy task. In this report we attempt to implement a machine learning approach for image segmentation and detection.We tried to learn and train mostly the random forest model to get an idea of the process which gave us an IoU value of 47.13%. The models previously used are Random Forest , CNN Random Forest Hybrid and U-NET giving an accuracy of above 50%.

INTRODUCTION

Europa, Jupiter's intriguing icy moon, defies expectations with a hidden ocean beneath its thick icy crust, sustained by Jupiter's tidal forces. Despite its desolate appearance, Europa's potential for life challenges conventional habitable zone theories. The red scars, known as Lineas, add mystery to its surface, formed by a blend of tidal forces and asynchronous rotation. The project aims to explore Lineas, unravel their mysteries, and analyse salts, hoping to discover organic compounds. The detection of hydrated salts fuels anticipation for potential organic findings on Europa, offering a glimpse into the broader possibilities of habitable icy moons beyond our solar system.

Machine learning methods could be utilized in the scientific exploration of celestial bodies such as Europa, one of Jupiter's moons. These techniques aid in the automated detection of distinctive surface features, such as "lineas". By analyzing image data with ML algorithms, researchers can uncover insights into Europa's geological characteristics and potential for sustaining life.

LITERATURE REVIEW

The papers studied, collectively demonstrate the versatility and effectiveness of machine learning (ML) techniques in image segmentation tasks across various domains. By leveraging techniques such as Convolutional Neural Networks (CNNs), Random Forests (RF), Conditional Random Fields (CRF), and Artificial Neural Networks (ANN), researchers have achieved notable success in segmenting images ranging from electron microscopy neuron segmentation to medical imaging and remote sensing. The comparison of different methods highlights the strengths and limitations of each approach, with some demonstrating superior performance in specific tasks while others provide viable alternatives, particularly in scenarios where GPU resources are limited. Overall, these studies underscore the importance of ML in tackling complex image segmentation challenges and offer valuable insights into the diverse applications of ML algorithms in image analysis.

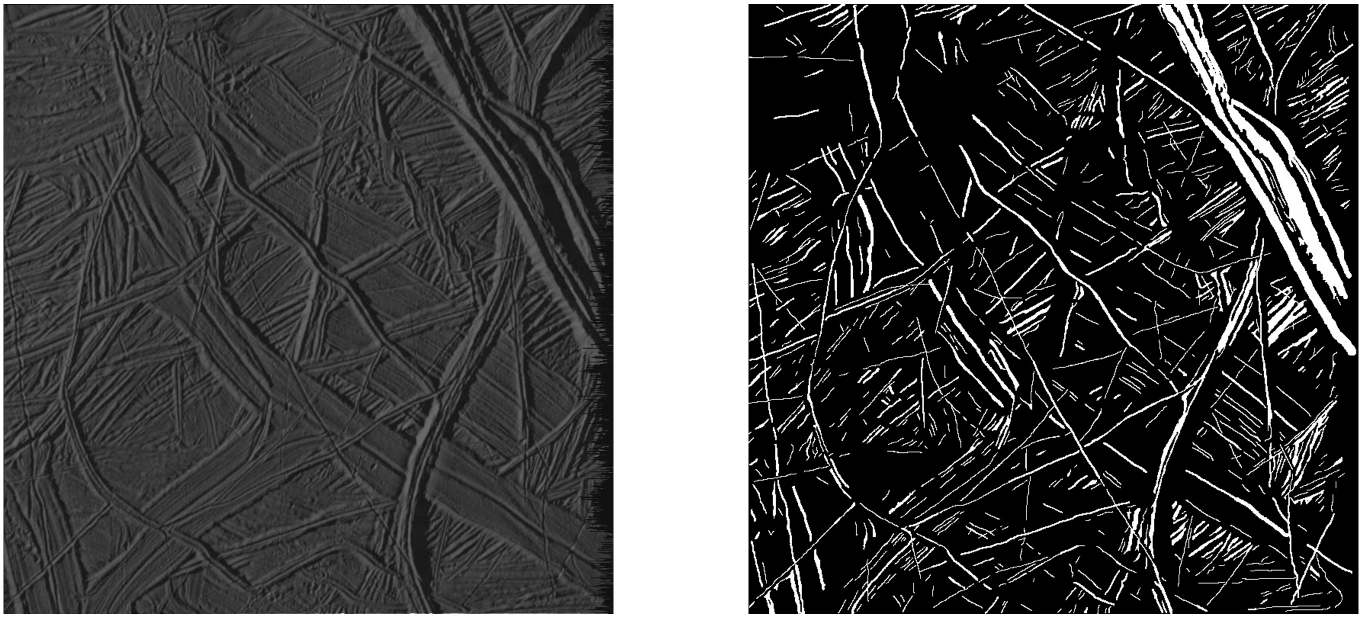
DATA COLLECTION AND PREPROCESSING

The images were sourced from Galileo's Solid State Imaging instrument and are publicly accessible through NASA via PDS-atlas.Seven 800×800 images were manually annotated by the previous project team.Then those images were each split into four parts each with a size of 400×400. Then the images and corresponding labels were organized into training and testing sets(70% and 30% correspondingly of the total dataset).In the dataset, the Lineas class, representing the minority, is significantly underrepresented in the dataset. Consequently, even if the

* pradeepkumar.baisakh@niser.ac.in

† anshuman.panda@niser.ac.in

model predominantly predicts pixels to belong to the majority class, it can achieve a high accuracy. To address this issue, we employed two techniques: Random Under Sampling (RUS), involving the random removal of majority class points, and Synthetic Minority Over-sampling Technique (SMOTE), which involves generating synthetic minority class datapoints. This approach notably enhances the Intersection over Union (IoU) metric.



Original image input

Masked image

FIG. 1

Feature Extraction

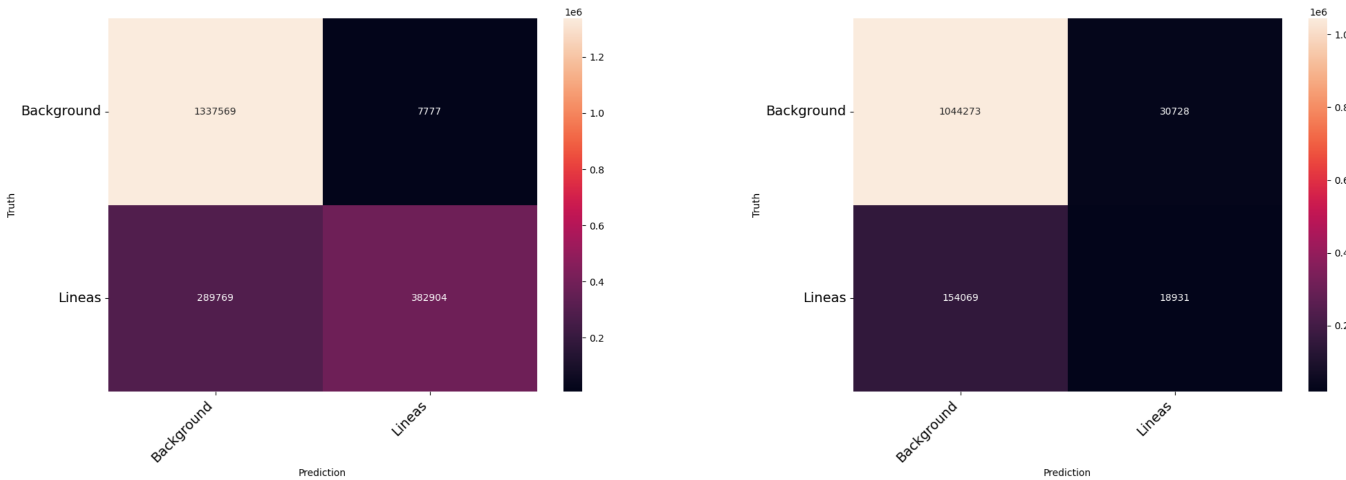
At first the images were converted to gray scale, grayscale conversion simplifies the subsequent processing steps while retaining essential information from the image.The original grayscale image is then flattened and added to the dataframe. This step transforms the 2D image matrix into a 1D array, where each element represents a pixel intensity value. Various image processing techniques were applied to extract features such as Gabor filters for texture features,Edge detection using Canny, Sobel, Roberts, Scharr, and Prewitt filters,Gaussian, Median, and Variance filters.Gabor filters helps with extracting texture features, some of the techniques like canny, sobel etc. help to capture edge related features, others help with pixel intensities and data normalization.Finally, all the extracted features are organized into a pandas DataFrame, where each row corresponds to an image and each column represents a specific feature. This structured representation facilitates further data manipulation and analysis. After extracting the features we got a total of 51 features and 4.16 million data points i.e. 4160000 rows x 51 columns matrix. Label Preparation was done by creating Masks corresponding to lineae and converting them to binary format (0 for no lineae, 1 for lineae). Also Binary masks were flattened for training.

	gabor24	Gaussian s7	gabor20	gabor4	gabor8	Gaussian s5	gabor12	gabor28	gabor16	gabor32	Gaussian s3	Median s5	Median s1	Median s7	Median s3	Original	Gaussian s1	Roberts	Variance s7	Variance s1
0	0.000000	0.247059	0.000000	0.000000	0.000000	0.254902	0.000000	0.000000	0.000000	0.000000	0.239216	0.105882	0.105882	0.105882	0.105882	0.058824	0.137255	0.000764	0.968627	0.968627
1	0.000000	0.247059	0.000000	0.000000	0.000000	0.250980	0.011765	0.011765	0.000000	0.000000	0.239216	0.113725	0.113725	0.113725	0.113725	0.105882	0.160784	0.000900	0.780392	0.780392
2	0.000000	0.247059	0.011765	0.011765	0.000000	0.250980	0.007843	0.007843	0.000000	0.000000	0.239216	0.113725	0.113725	0.113725	0.113725	0.113725	0.180392	0.001014	0.309804	0.309804
3	0.000000	0.243137	0.000000	0.000000	0.000000	0.247059	0.000000	0.000000	0.003922	0.003922	0.235294	0.113725	0.113725	0.113725	0.113725	0.078431	0.172549	0.000982	0.643137	0.643137
4	0.011765	0.243137	0.000000	0.000000	0.011765	0.243137	0.000000	0.000000	0.007843	0.007843	0.231373	0.090196	0.090196	0.090196	0.090196	0.066667	0.156863	0.000838	0.050980	0.050980

FIG. 2: Part of Dataset with 20 top features

MODEL EVALUATION

Our model was evaluated with some values like pixel accuracy, where the model gets a score of 1 for each correct pixel classification (linea or surface), and 0 for misclassifications, then the accuracy is the total sum of these values (1 or 0) over total number of pixels. IoU(Intersection over Union) value is the intersection of the predicted image and the mask divided by the unionof the predicted image and the mask. The confuion matrix was also obtained.



Training data confusion matrix

Test data confusion matrix

FIG. 3

RESULTS

The model achieved a mean Intersection over Union (IoU) of 69.04% on the training data, indicating a good performance in accurately segmenting Lineas. This metric measures the overlap between the predicted Lineas and the ground truth masks, providing insight into the model's segmentation accuracy. However, the IoU dropped to 47.13% on the testing data, suggesting some degree of overfitting to the training set. The pixel accuracy, which measures the overall correctness of pixel-level classifications, remained high at 85.26% for training and 38.12% for testing. These results indicate that while the model performs well on the training data, it struggles to generalize to unseen examples, warranting further investigation into model generalization and potential avenues for improvement.

Data Type	mIoU	IoU	Pixel accuracy
Training	69.04%	56.27%	85.26%
Testing	47.13%	9.29%	38.12%

CONCLUSION

In conclusion, by trying to rebuild this machine learning model for detecting Lineas on Europa's surface helped us learning the process of data processing, feature extraction, model training and about model performances. Our method of employing feature extraction techniques and training the model with Random Forest, the achieved accuracies show that this model now is not practically applicable for scientific use. The evaluation metrics, including pixel accuracy and IoU, demonstrate the model's effectiveness in detecting Lineas. We are planning to build other models using different ML algorithms or some hybrid algorithms which could give us better predictions while detecting the lineas, paving the way for deeper insights into Europa's intriguing geology.

[1] P. Sathya and M. Lakshmanan, Classification and segmentation in satellite imagery using back propagation algorithm of ann and k-means algorithm, International Journal of Machine Learning and Computing , 422 (2011).

[2] N. Plath, M. Toussaint, and S. Nakajima, Multi-class image segmentation using conditional random fields and global classification, in *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09 (Association for Computing Machinery, New York, NY, USA, 2009) p. 817–824.

[3] D. Hartmann, D. Müller, I. Soto-Rey, and F. Kramer, Assessing the role of random forests in medical image segmentation (2021), arXiv:2103.16492 [eess.IV].

[4] G. Cao, S. Wang, B. Wei, Y. Yin, and G. Yang, A hybrid cnn-rf method for electron microscopy images segmentation, Journal of Biomimetics, Biomaterials, and Tissue Engineering **18** (2013).

[5] N. P. Das, R. G. Nazareth, S. Mishra, and G. S. Thangjam, Automated Detection of Lineaments on the Surface of Europa Using Machine Learning, in *53rd Lunar and Planetary Science Conference*, LPI Contributions, Vol. 2678 (2022) p. 2709.