Ocean Pollution Detection

Contents

Introduction

The world's oceans face growing threats from various forms of pollution, including plastic waste, oil spills, and chemical runoff. Understanding the nature and impact of ocean pollution is a critical first step in addressing this global challenge.

Importance of Detecting Ocean Pollution

1 Ecosystem Health

Detecting and monitoring ocean pollution is crucial for preserving marine ecosystems and protecting vulnerable wildlife.

2 Human Impact

Ocean pollution can have far-reaching consequences for human communities that depend on the sea for food, livelihood, and recreation.

3 Policy and Regulation

Accurate data on ocean pollution levels can inform policy decisions and drive the development of effective regulations and cleanup efforts.

Image Preprocessing

1 **Colour Correction:**

Colour correction employs a white balancing process to remove unwanted colour casts, restoring the original colour integrity of the images. This enhances the overall quality, making them clearer and easier to analyze.

2 **Noise Reduction:**

Underwater images frequently suffer from noise, which can distort visual clarity. These methods suppress unwanted noise in the images, resulting in cleaner and visually appealing outputs. By reducing noise, we enhance image quality, facilitating more accurate content analysis.

3 **Sharpness Enhancement:**

To further enhance the quality and clarity of underwater images, we apply sharpness enhancement techniques. These methods improve the definition and crispness of edges and details in the images, making them sharper and visually striking. By enhancing sharpness, we highlight important features and elements, aiding in the identification and assessment of pollution elements such as plastic and metals. .

Why CNN?

1. Chose CNN for feature extraction due to advantages over PNN and SVM:
   - PNNs are memory-intensive and sensitive to architecture/hyperparameters.
   - SVMs lack probabilistic outputs and interpretability, making CNNs more suitable.
2. CNNs automatically learn hierarchical representations, ideal for tasks like image recognition.
3. Scalable and adept at handling large datasets, crucial for real-world applications.

More about CNN classifier:

1. Trained CNN classifier for image categorization into plastic, metals, and other pollutants.
2. CNN includes convolutional, pooling, and fully connected layers for image classification.
3. Flatten Layer: Converts output from convolutional layers into a one-dimensional vector for dense layers.
4. Dense Layers: 128-unit layer with ReLU activation for learning complex patterns, and a 3-unit layer with softmax activation for multi-class classification.
5. Softmax activation ensures probability distribution across classes, aiding accurate predictions.

Database Analysis

1. Utilized Kaggle dataset comprising over 2560 images of varying resolutions (640x480 to 1200x900 pixels).
2. Standardized image size to 224x224 pixels for compatibility with pre-trained models, efficient processing, and optimal classifier performance.
3. Standardization of resolution ensures representative and comprehensive training/testing of the classifier on the diverse Kaggle dataset.

Algorithm

Convolutional Layers:

- First layer: Input images of 224x224 pixels with 3 channels(RGB), 6 filters of 5x5 size, ReLU activation for low-level features.
- Second layer: 16 filters of 5x5 size with ReLU activation for deeper feature extraction.
- Third layer: 64 filters of 3x3 size with ReLU activation for capturing complex features.

Max Pooling Layers :

- After each convolutional layer, a maxpooling layer with a pool size of 2x2 and stride of 2x2 reduces spatial dimensions and retains important features.

Flatten Layer:

- Converts 3D feature maps from conv layers into 1D feature vector for fully connected layers.

Dense Layers:

- First dense layer: 128 neurons with ReLU activation for high-level feature learning, includes dropout regularization with a rate of 0.5 to prevent overfitting.
- Output Layer: Final dense layer with 3 neurons and softmax activation for classification task, producing class probabilities.

Useful Links:
https://deeplizard.com/resource/pavq7noze3
https://deeplizard.com/resource/pavq7noze2
https://adamharley.com/nn_vis/cnn/3d.html

```
Model: "sequential_5"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d_15 (Conv2D)          (None, 224, 224, 6)       456

 max_pooling2d_15 (MaxPoolin  (None, 112, 112, 6)      0
 g2D)

 conv2d_16 (Conv2D)          (None, 112, 112, 16)      2416

 activation_5 (Activation)   (None, 112, 112, 16)      0

 max_pooling2d_16 (MaxPoolin  (None, 56, 56, 16)       0
 g2D)

 conv2d_17 (Conv2D)          (None, 54, 54, 64)        9280

 max_pooling2d_17 (MaxPoolin  (None, 27, 27, 64)       0
 g2D)

 flatten_5 (Flatten)         (None, 46656)             0

 dense_10 (Dense)            (None, 128)               5972096

 dropout_5 (Dropout)         (None, 128)               0

 dense_11 (Dense)            (None, 3)                 387

=================================================================
Total params: 5,984,635
Trainable params: 5,984,635
Non-trainable params: 0
_________________________________________________________________
```

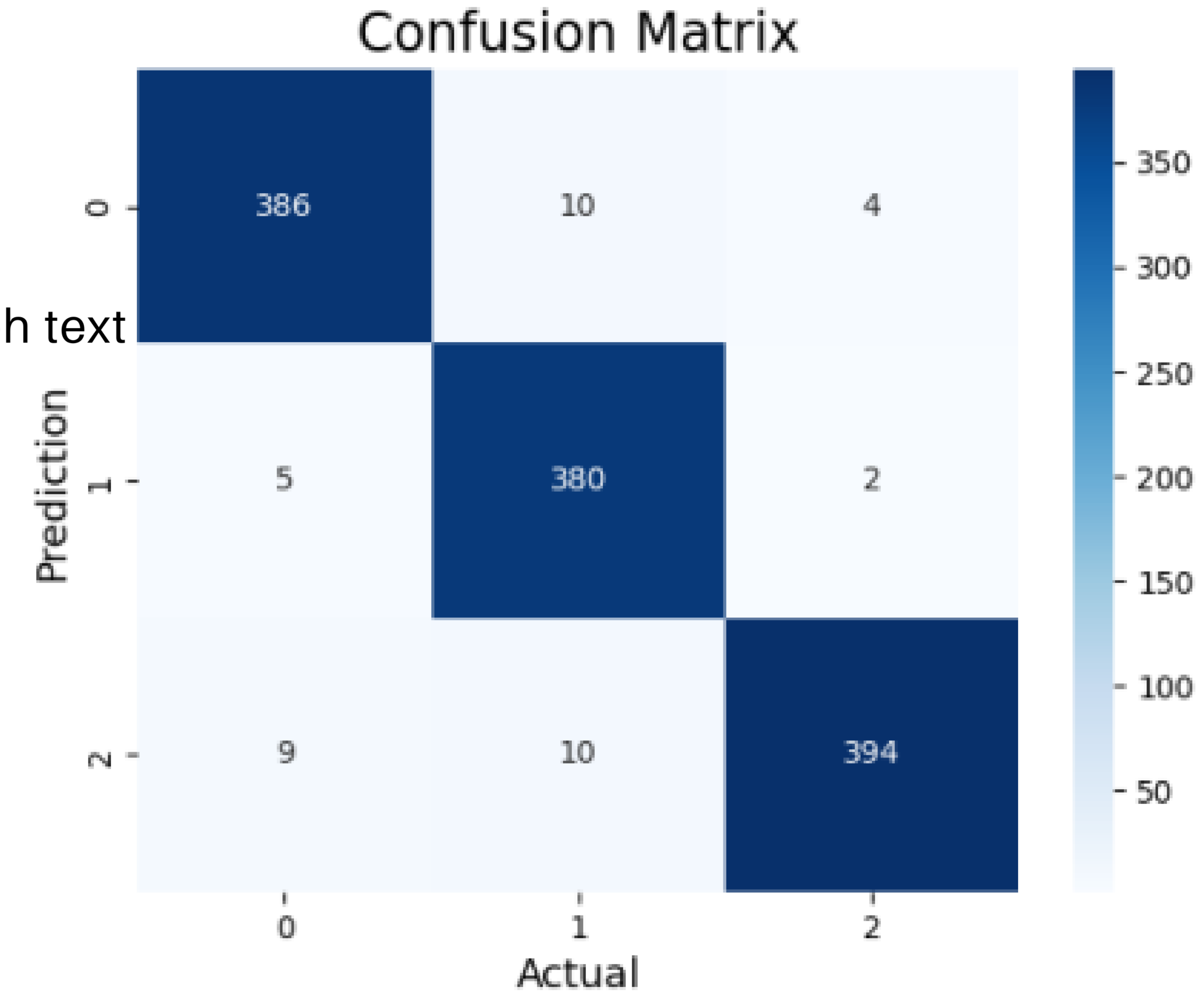
Simulation results of the proposed approach

Numerical Results

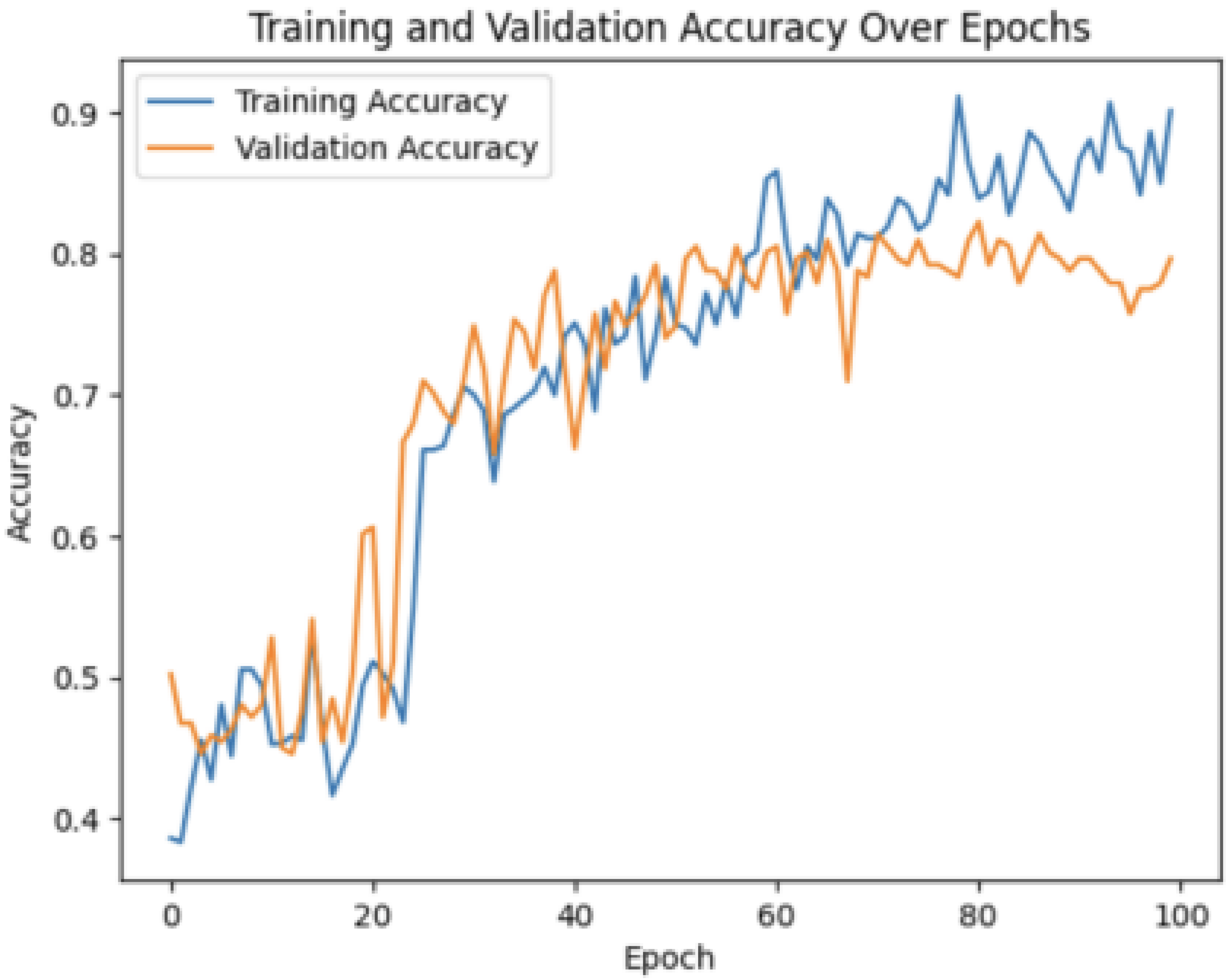
From the matrix, we can see that the classifier can identify images correctly with the help of the diagonal value. The image preprocessing has increased the accuracy of the obtained output. The classification done with the help of the CNN model is also very accurate, as shown in the graph above.

Your paragraph text

Accuracy	96.67
Average Recall	96.66
Average Precision	97.01



0 - Non-Plastic, 1 - Plastic, 2 - Metal



Plotting epoch on the horizontal axis and accuracy after each epoch on the vertical axis.

Conclusion

1. Utilising technological milestones
2. Efficiency and optimisation
3. Technological Responsibility

Our approach provides prompt data for environmental surveillance and control, intending to lessen the impact of human actions on marine habitats. Through teamwork across various fields and imaginative technologies, we anticipate a future where successful pollution identification results in improved oceanic conditions

Contributions

Rishi Arora (202101025)	Literature survey, ideation and data collection
Mehul Rawal (202101033)	Proposed Approach and data collection
Parth Parmar(202101077)	algorithm designing, coding and debugging
Adhya Desai (202101111	Presentation making, literature survey and ideation
Aum Patel (202101118)	generating and analyzing results
Hirmi Patel(202101151)	data collection, ideation report writing, learnt CNN and helped in coding the algorithm
Het Patel (202101162)	Proposed approach and coding of preprocessing
Kirtan Soni (202101197)	Worked on whole coding program
Aryan Vagh (202101221)	ideation, data collection, NGO idea research, report writing in latex
Vansh Joshi(202101445)	learnt ML modelling, assisted in algorithm designing and coding and latex writing

Thank you!