# Classification of Fish Species Using Convolutional Neural Networks (report)

## Introduction

The aim of this project is to classify fish species based on image data using a Convolutional Neural Network (CNN). Accurate identification of fish species is essential for applications such as biodiversity conservation, fisheries management, and ecological research. This report outlines the process of building, training, and evaluating a CNN model for this purpose.

## Data Collection and Preprocessing

The dataset comprises 3,960 images from 468 fish species, captured under three different conditions: controlled, out-of-the-water, and in-situ. The images were pre-processed and resized to a common size of 224x224 pixels. The labels were encoded using one-hot encoding.

Steps:

**Image Loading:** Images were loaded from the dataset directory.

**Image Resizing:** All images were resized to 224x224 pixels.

**Label Encoding:** Labels were encoded into numerical format and then converted to categorical data.

**Data Splitting:** The dataset was split into training (80%) and test (20%) sets.

**Normalization:** Pixel values were normalized to the range [0, 1].

## Model Building

A Convolutional Neural Network (CNN) was built using TensorFlow and Keras. The architecture included convolutional layers, max-pooling layers, a flatten layer, dense layers, and a dropout layer to prevent overfitting.

### Model Architecture:

**Convolutional Layer:** 32 filters, kernel size 3x3, ReLU activation

**Max-Pooling Layer:** Pool size 2x2

**Convolutional Layer:** 64 filters, kernel size 3x3, ReLU activation

**Max-Pooling Layer:** Pool size 2x2

**Convolutional Layer:** 128 filters, kernel size 3x3, ReLU activation

**Max-Pooling Layer:** Pool size 2x2

**Flatten Layer**

**Dense Layer:** 128 units, ReLU activation

**Dropout Layer:** 0.5 dropout rate

**Dense Layer:** Output layer with softmax activation

## Model Training

The model was trained on the training dataset using the Adam optimizer, categorical cross-entropy loss function, and accuracy as the performance metric. The training process was monitored using a validation dataset.

### Training Parameters:

**Epochs:** 20

**Batch Size:** 32

**Optimizer:** Adam

**Loss Function:** Categorical Cross-Entropy

## Model Evaluation

The trained model was evaluated on the test dataset, achieving a test accuracy of 80.69%. This indicates that the model is effective in classifying fish species based on the provided image data.

### Evaluation Metrics:

**Accuracy:** 80.69%

A graph of different colored lines

Description automatically generated

Figure 1, Training and Validation Plots

**Loss:** Monitored during training and validation phases

A screenshot of a computer

Description automatically generated

Figure , Monitored training and validation

## Conclusion

This study successfully developed a CNN model capable of classifying fish species from image data with an accuracy of 80.69%. This model can be further refined and optimized for even higher accuracy and robustness, potentially aiding in environmental monitoring, fisheries management, and ecological research.

## Future Work

Future improvements can include:

Experimentation with more complex model architectures

Hyperparameter tuning for better performance

Transfer learning using pre-trained models