Fish Species Classification Using Deep Learning

# 1. Introduction

Fish species classification plays a vital role in the fishing industry, aquaculture, and biodiversity research. Accurate identification helps in sustainable fishing, quality control in seafood markets, and ecological balance. In this project, deep learning models were applied to classify fish species from images.

# 2. Dataset Description

The dataset consists of images belonging to 11 fish categories:  
['animal fish', 'animal fish bass', 'fish sea\_food black\_sea\_sprat', 'fish sea\_food gilt\_head\_bream', 'fish sea\_food hourse\_mackerel', 'fish sea\_food red\_mullet', 'fish sea\_food red\_sea\_bream', 'fish sea\_food sea\_bass', 'fish sea\_food shrimp', 'fish sea\_food striped\_red\_mullet', 'fish sea\_food trout']  
  
Data was split into training, validation, and testing sets. Preprocessing was carried out using ImageDataGenerator with rescaling and augmentation such as rotation, shift, zoom, and flipping.

# 3. Methodology

Two approaches were explored:  
1. CNN from Scratch – A custom CNN model with convolutional, pooling, dropout, and dense layers.  
2. Transfer Learning – Pre-trained models (VGG16, ResNet50, MobileNet, InceptionV3, EfficientNetB0) were fine-tuned for fish classification. The base models were frozen initially and the final layers were retrained with a lower learning rate for better generalization.

# 4. Implementation Flow

The implementation involved the following steps:  
- Data preprocessing using ImageDataGenerator.  
- Model building (CNN and Transfer Learning models).  
- Training with callbacks such as EarlyStopping and ModelCheckpoint.  
- Evaluation on validation and test sets.  
- Comparison of models based on accuracy, precision, recall, and F1-score.  
- Best model selection.  
- Deployment using a Streamlit application.

# 5. Results & Evaluation

All models were evaluated on the validation and test sets. Key results are summarized below:  
  
Validation Results:  
- VGG16: Accuracy ~10%  
- ResNet50: Accuracy ~9%  
- MobileNet: Accuracy ~9%  
- InceptionV3: Accuracy ~8%  
- EfficientNetB0: Accuracy ~10%  
- CNN (scratch): Accuracy ~10%  
  
Test Results (after fine-tuning):  
- VGG16: 99.2%  
- ResNet50: 71.8%  
- MobileNet: 99.8%  
- InceptionV3: 99.8%  
- EfficientNetB0: 23%  
- CNN (scratch): 90.4%  
  
MobileNet and InceptionV3 performed the best with ~99.8% accuracy. EfficientNetB0 underperformed significantly, while ResNet50 struggled with certain classes. Small classes like 'animal fish bass' were harder to classify correctly.

# 6. Discussion

MobileNet and InceptionV3 achieved the highest accuracy, demonstrating the strength of transfer learning on image datasets. CNN from scratch achieved reasonable accuracy (~90%) but lagged behind the fine-tuned models. EfficientNetB0 failed likely due to insufficient fine-tuning or dataset size. The results highlight the importance of proper fine-tuning and balanced data.  
  
Challenges included class imbalance (e.g., very few 'animal fish bass' samples) and misclassification across visually similar fish categories.

# 7. Deployment

A Streamlit application was developed for deployment. Users can upload a fish image, and the model predicts the fish category along with confidence scores. To handle non-fish images, a confidence threshold was implemented. Images with prediction confidence below 60% are labeled as 'Unknown / Not a Fish'. This makes the application more robust in real-world use.

# 8. Conclusion & Future Work

This project demonstrated the application of deep learning in fish species classification. Transfer learning models like MobileNet and InceptionV3 provided state-of-the-art results, achieving ~99.8% accuracy. Future improvements include:  
- Expanding the dataset and balancing underrepresented classes.  
- Adding an 'Other/Background' class to better handle non-fish images.  
- Applying hyperparameter tuning for further optimization.  
- Deploying the model as a cloud service for wider accessibility.