UMBC

Enhancing Classification of Aquatic Species through Supervised Contrastive Learning and Advanced Image Super-Resolution

Presented By:

Sadia Nasrin Tisha

Geethanjali Nallani

Outline

- Problem Statement
- Dataset Description
- Super Resolution Techniques
- Classification Task
- Result
- Discussion
- Limitations
- Future Works

Problem Statement

•**Background**:

- Invasive species like Zebra and Quagga mussels have become ecological threats in North America.

- US government is funding the QZAP program, which is spending roughly $2M / year.

•**Impact**:
These species damage ecosystems and infrastructure, causing substantial economic losses.

•**Current Detection Methods**:
Conventional methods are costly, time consuming, and require expert knowledge.

•**Challenges with Traditional Methods**:
Reliance on manual sampling and microscopy.

•**Need for Automation**:
Importance of developing automated monitoring and efficient methods for early detection of invasive species

Figure 1: Boat Propeller clogged by Zebra mussel

Invasive Species Larvae Dataset

- **Source**: High-definition video recordings of aquatic streams from the Colorado River near Davis Dam, Arizona.
- **Processing Method**: Kalman Filter-based proprietary algorithm for identifying, tracking, and extracting larvae imagery.
- **Dataset Composition**:
    - Total Organisms: 6905 (1220 organisms that are invasive, 5685 non-invasive).
    - Total Images: 88050 (44646 images of invasive species, 43404 images of non-invasive species).

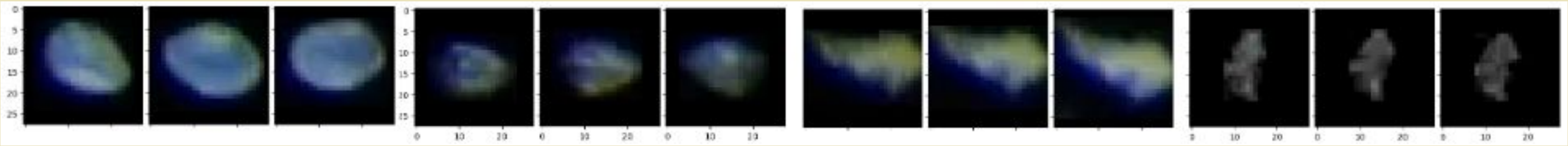


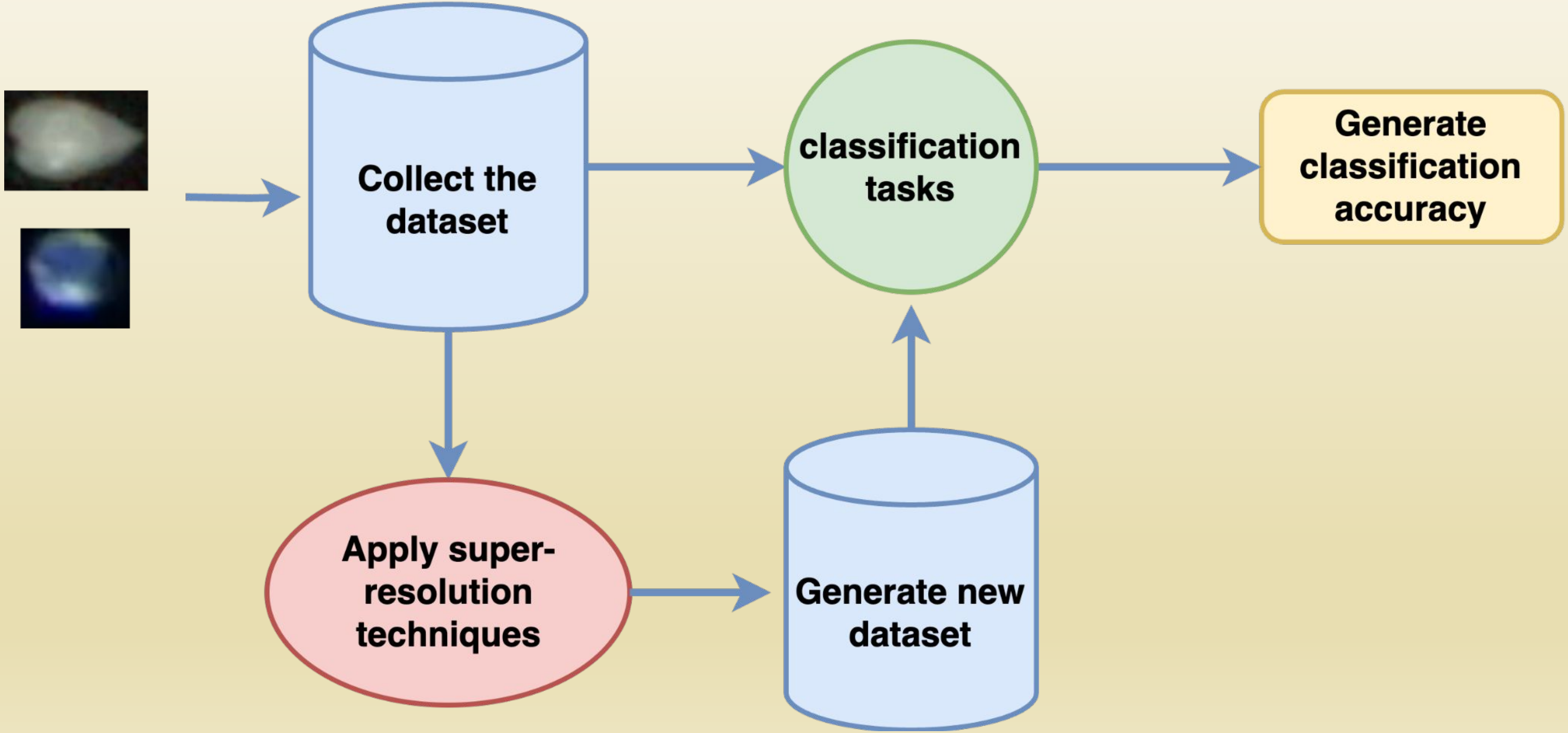
Figure 2: Example of invasive dreissenid and non-invasive species larvae

Invasive Species

Non Invasive Species

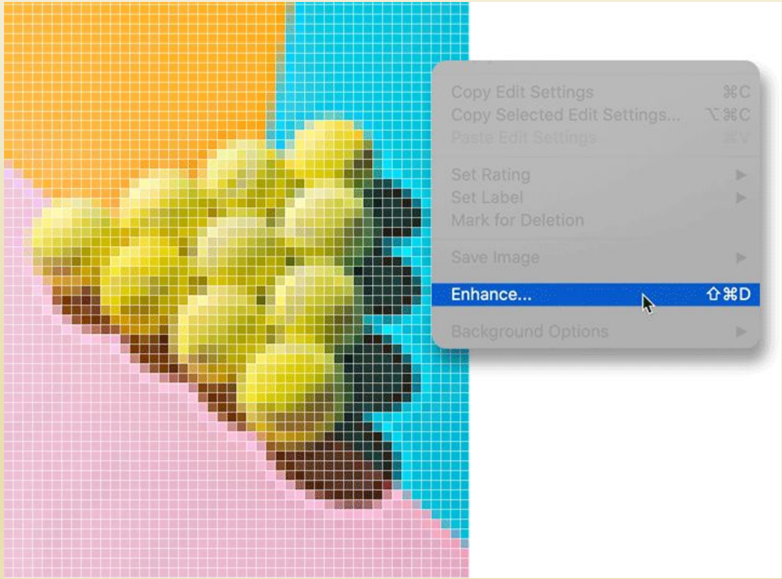
How does the enhancement of image quality through super-resolution techniques affect the performance of supervised contrastive learning models in the classification of aquatic species?

Workflow

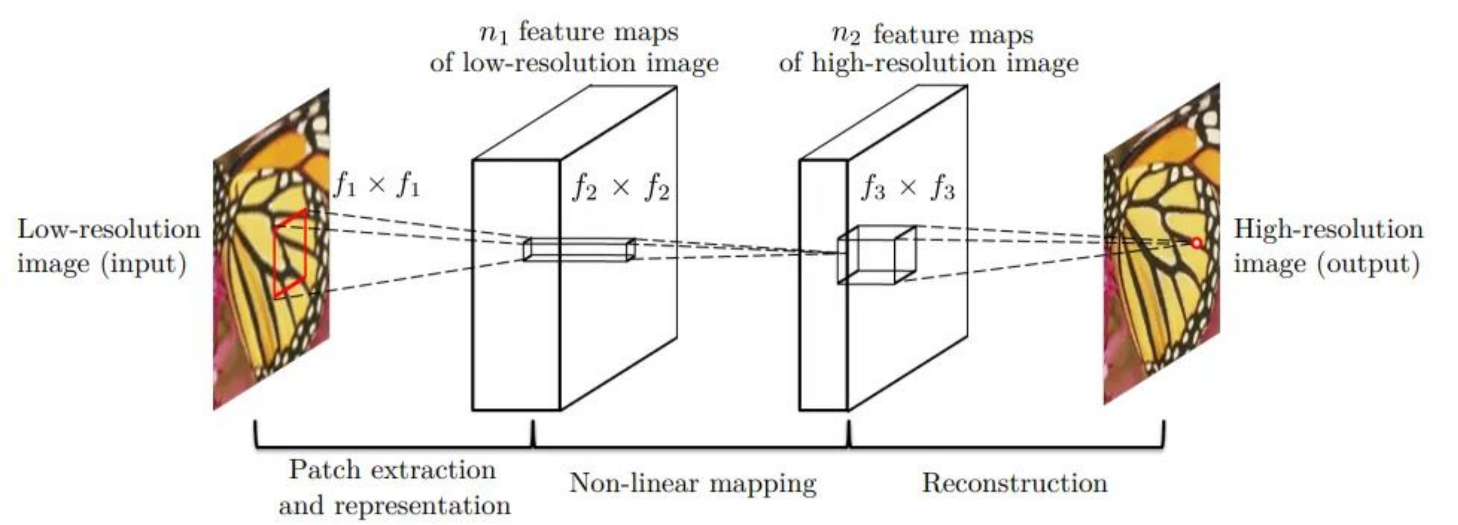


Collect the dataset
classification tasks
Generate classification accuracy
Apply super-resolution techniques
Generate new dataset

Super Resolution Techniques



SRCNN ARCHITECTURE



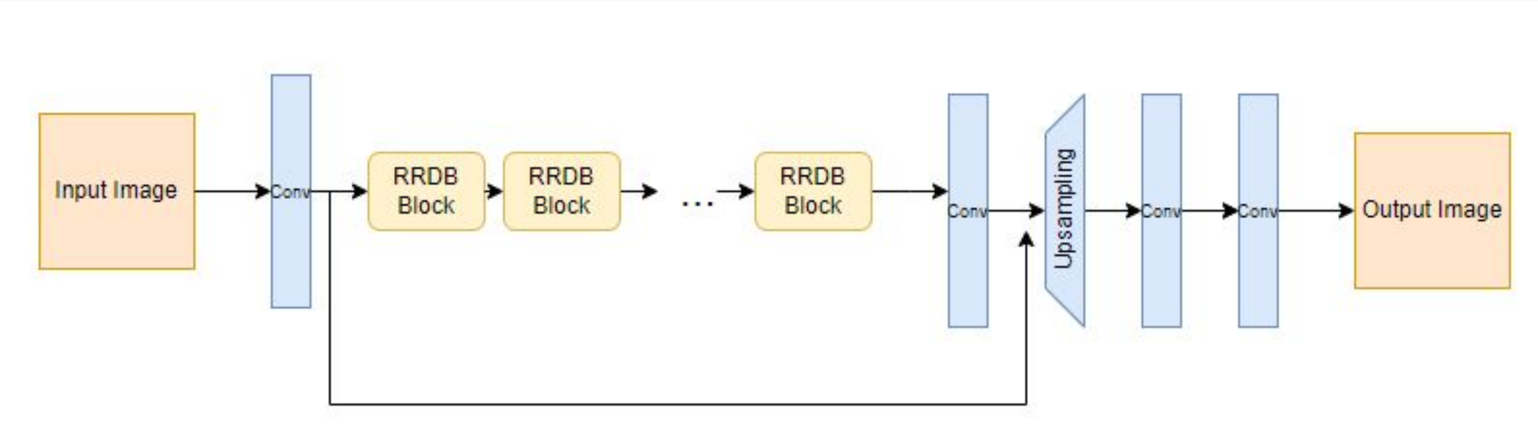
Ref: https://medium.com/analytics-vidhya/srcnn-paper-summary-implementation-ad5cea22a90e

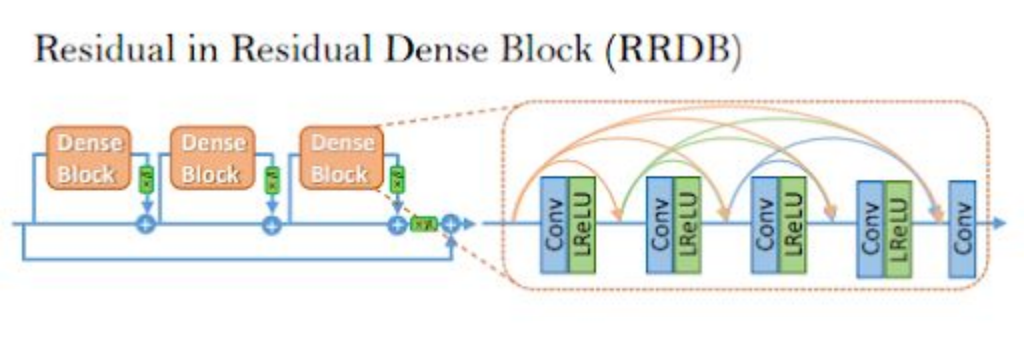
Output of SRCNN

SRCNN Model:


Original
Degraded
SRCNN Reconstructed
PSNR: 34.46
MSE: 69.85
SSIM: 0.924
PSNR: 34.73
MSE: 65.69
SSIM: 0.933

ESRGAN ARCHITECTURE



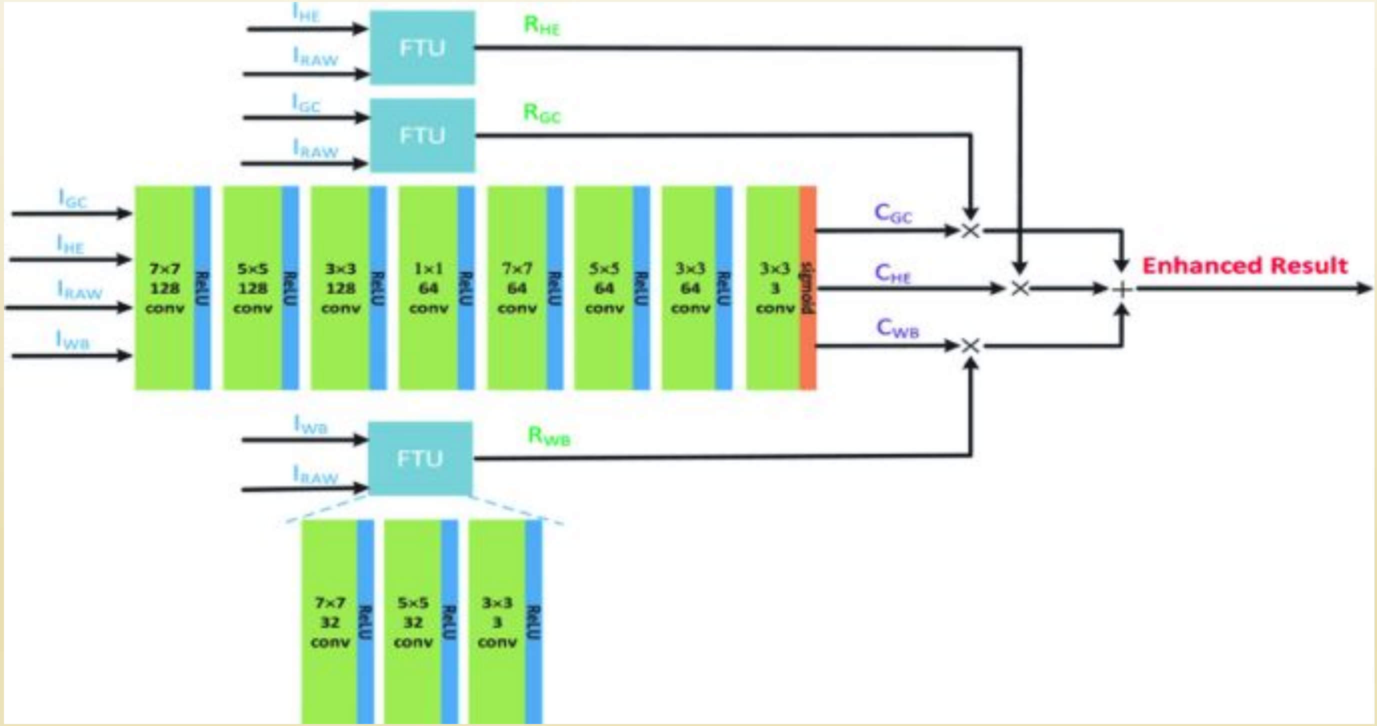


Ref: https://www.tensorflow.org/hub/tutorials/image_enhancing

Output of ESRGAN

ESRGAN Model:

WATER-NET ARCHITECTURE

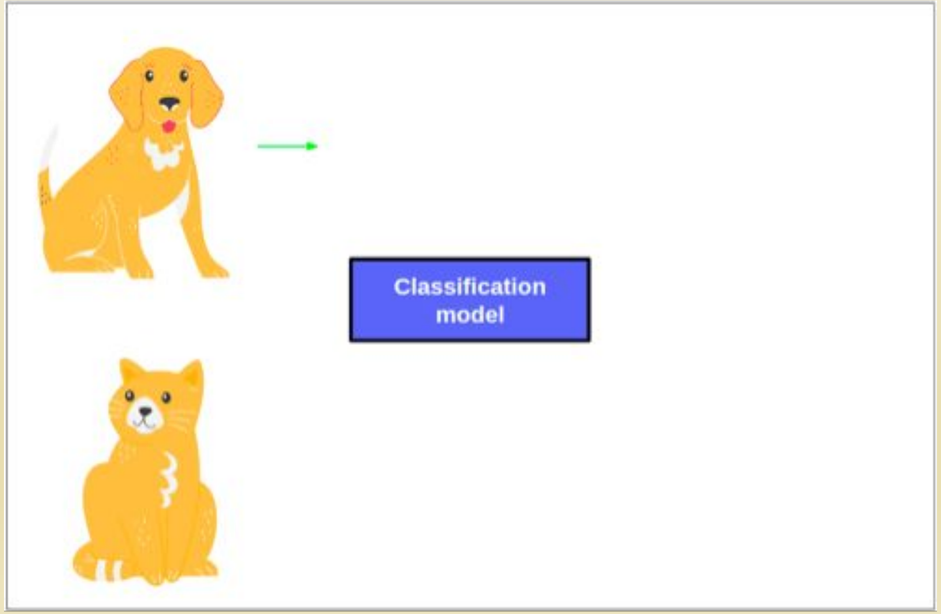


ref:https://arxiv.org/pdf/1901.05495

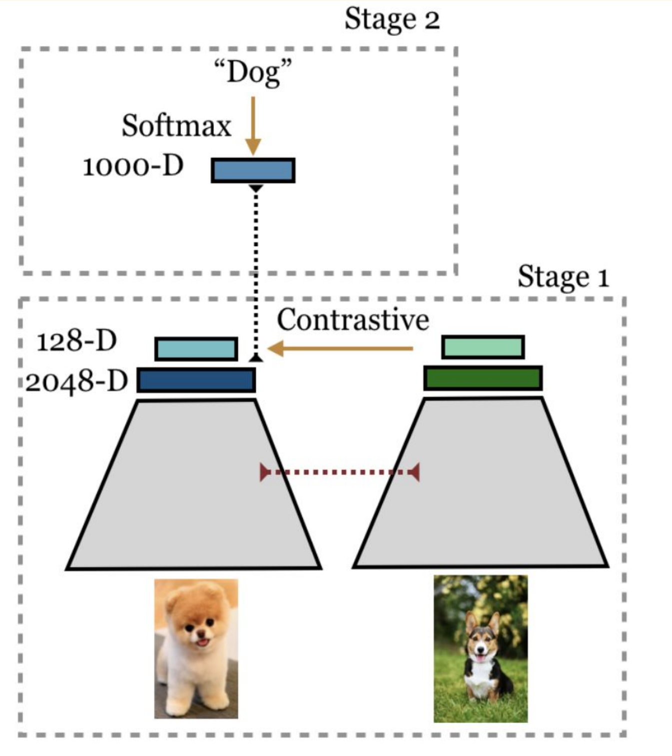
Output of WATER-NET MODEL

Water-Net Model:

Classification task



Supervised Contrastive Learning



Ref: Khosla, Prannay, et al. "Supervised contrastive learning." Advances in neural information processing systems 33 (2020): 18661-18673.

Image Augmentation Methods

Normalization

This layer standardizes the images to have a mean of 0 and a standard deviation of 1.

- Help neural networks train faster and more effectively
- Ensures that the input data varies within a similar range.

Random Flip

This layer randomly flips the images horizontally. Horizontal flipping is a simple way to increase the diversity of the training dataset without collecting new images.

- Help in improving the robustness of the model
- Making it better at generalizing from the training data to new, unseen data.

Random Rotation

This layer randomly rotates the images by up to 0.02 radians (approximately 1.15 degrees).

- Useful for training models that need to recognize objects in different orientations.

- **Baseline classifier** is trained the on encoder and the classifier parts are trained together as a single model to minimize the cross entropy loss.

- The **supervised contrastive model** is trained in two phases:
  In the first phase, the encoder is pre-trained to optimize the supervised contrastive loss
  In the second phase, the classifier is trained using the trained encoder with its weights freezed; only the weights of fully-connected layers with the softmax are optimized.

- **Convolutional Neural Network(CNN)** model consists of two convolutional layers followed by max pooling, a flatten layer, and two dense layers with the final output layer using softmax activation for binary classification. The model is compiled with the Adam optimizer and categorical cross-entropy loss function.

Result

Model	Standard resolution	SRCNN	ESRGAN	Water-Net
Baseline classification model	**96.29%**	91.34%	**95.28%**	88.71%
Supervised contrastive learning	89.43%	87.63%	**96.96%**	88.19%
Convolutional Neural network	86.00%	81.56%	**91.06%**	89.39%

Discussion

- **ESRGAN** showed an improvement in classification accuracy over standard resolution images.
- This suggests that ESRGAN's method of enhancing texture and detail particularly beneficial for the features necessary in species classification.

- **SRCNN and WaterNet**, despite improving certain image quality metrics, did not lead to better classification performance and in some cases, reduced accuracy.
- This might be attributed to the types of image alterations these techniques introduce, such as noise or loss of critical detail.

Limitations

- Training Data Compatibility
- Sensitivity to Super Resolution Artifacts
- Model Sensitivity to Image Quality Changes

Future Works

- Enhance the image quality
- Enhanced Model Training
- Use of Data Augmentation
- Hybrid SR Techniques
- Cross-Domain Validation
- Exploring Alternative Machine Learning Approaches and Dataset


Thank You
Any question?