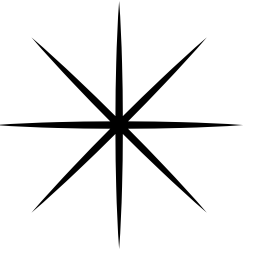


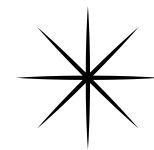


Ilaria Tangorre 5307483 - Greta Riva 5306762



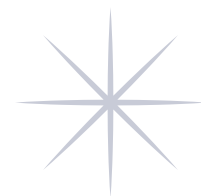
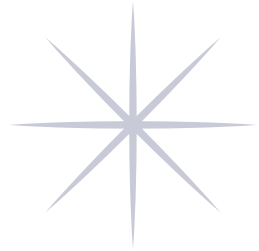
Diamond Carat Regression

Deep Learning for AI Project

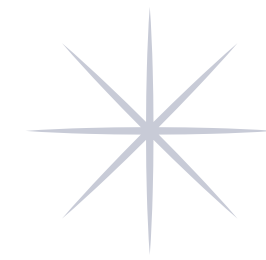


Project goal

Develop a model capable of predicting diamond carat values from images using a deep neural network



Data exploration



Subfolders



Diamond_data.csv



Composed by 14 variables, but for our project we are only interested in:

- **path_to_img:** it contains the URLs of the images not divided by classes
- **carat:** quantitative variable, with a minimum value of 0.5 and a maximum value of 8



48765 elements

Carat: 1.52



Carat: 0.9



Carat: 0.7



Carat: 0.7



cushion



heart



princess



round



oval



marquise



pear



emerald



Data preprocessing

1. Data cleaning

- Dataset was filtered to include only diamonds with carats ≤ 2.2 , enhancing regression reliability
- Rows with invalid stock numbers were removed

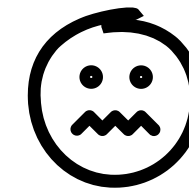
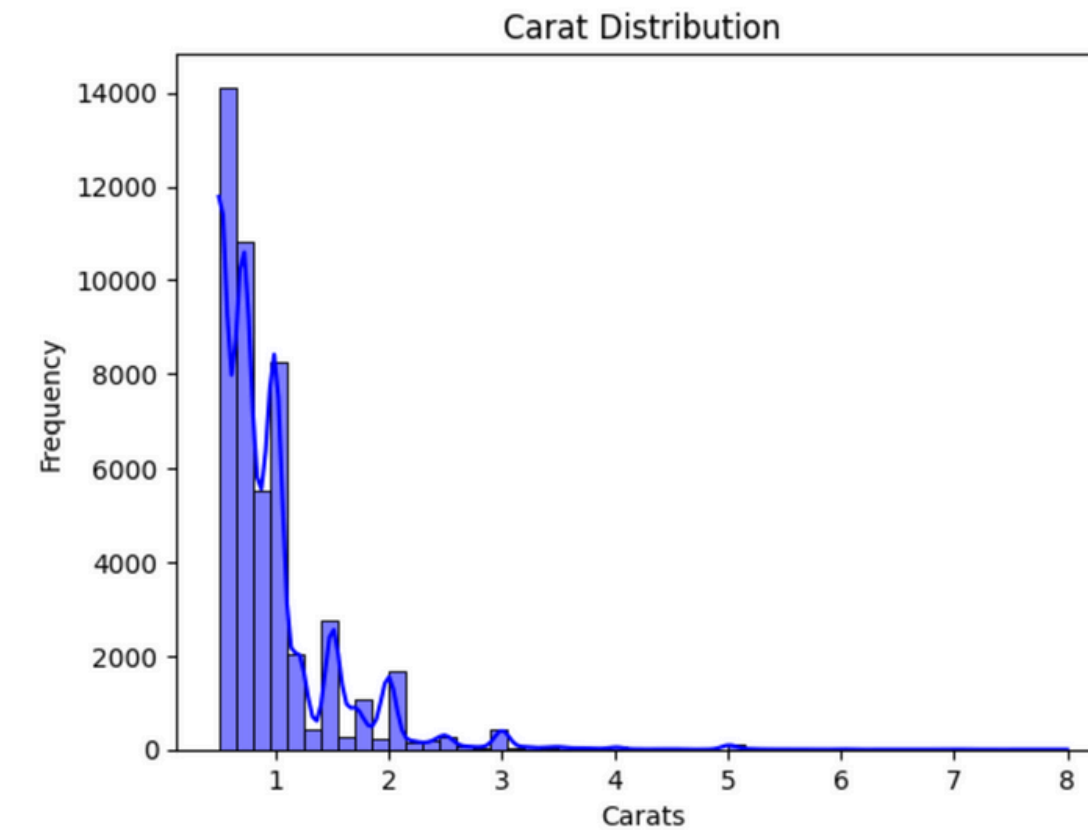
2. Data augmentation

- resize (224x224)
- image normalization

3. Carat scaling

it maps the carat values in a 0-1 range

$$X_{\text{scaled}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$



NOT BALANCED!

Why GoogLeNet?

- **Ability to capture multiple scales:** GoogLeNet, with its "Inception blocks," extracts features at different resolutions, ideal for diamond images with complex variations.
- **Computational efficiency:** Optimized to handle large datasets (48,765 images) without compromising performance.
- **Adaptation to regression:** Although designed for classification, its ability to extract complex features makes it suitable for predicting diamond carats.

Model implementation

MODEL

b1: 7X7 Convolution + ReLU + 3x3 MaxPool.

b2: 1x1 Convolution + ReLU + 3x3 Convolution + 3x3 MaxPool.

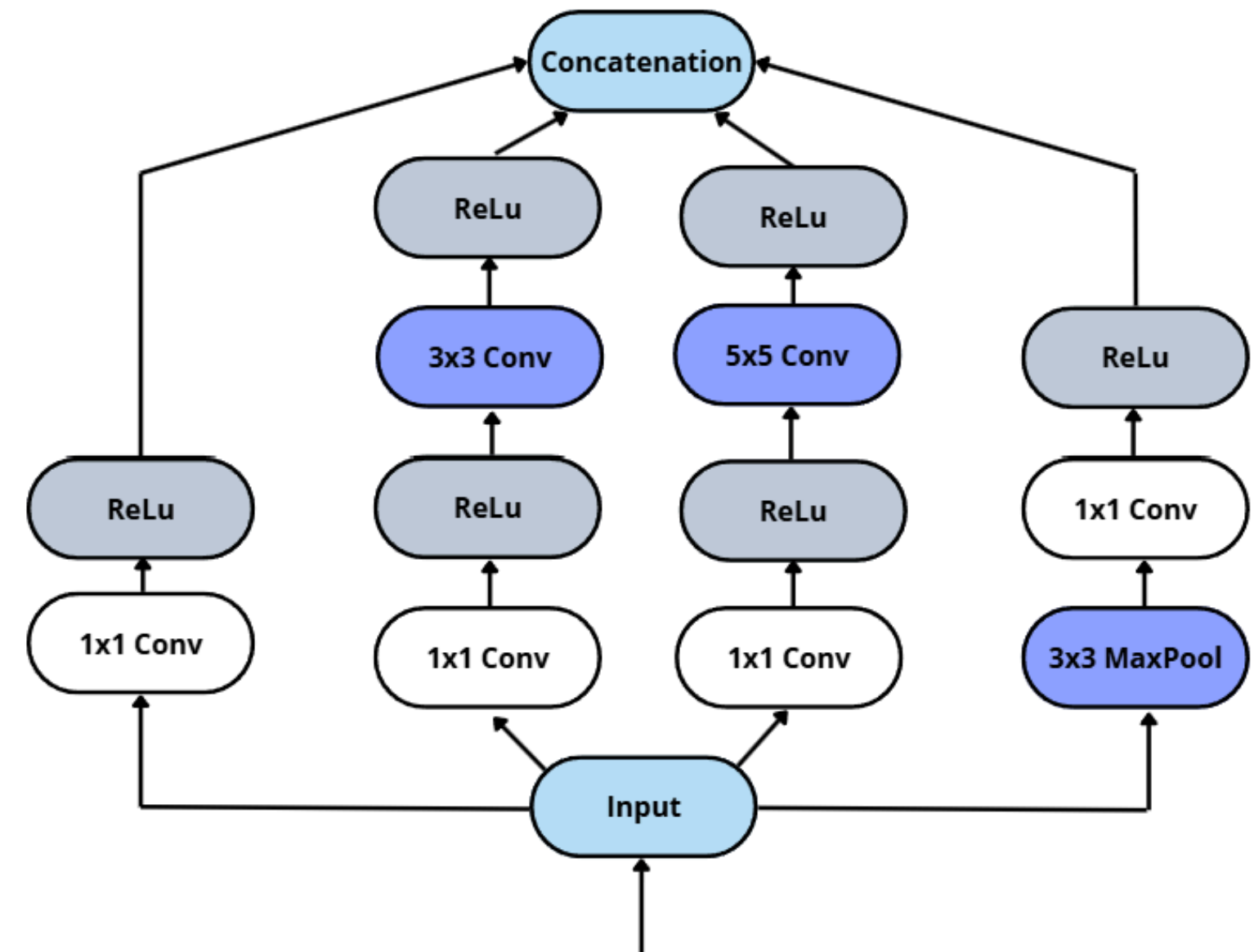
b3: Inception block + Inception block + 3x3 MaxPool.

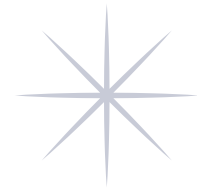
b4: 5 Inception blocks with multiple convolutions at different scales + 3x3 MaxPool.

b5: Inception block + inception block + Dropout ($p=0.4$) + 1x1 Adaptive Average Pooling + Flatten.

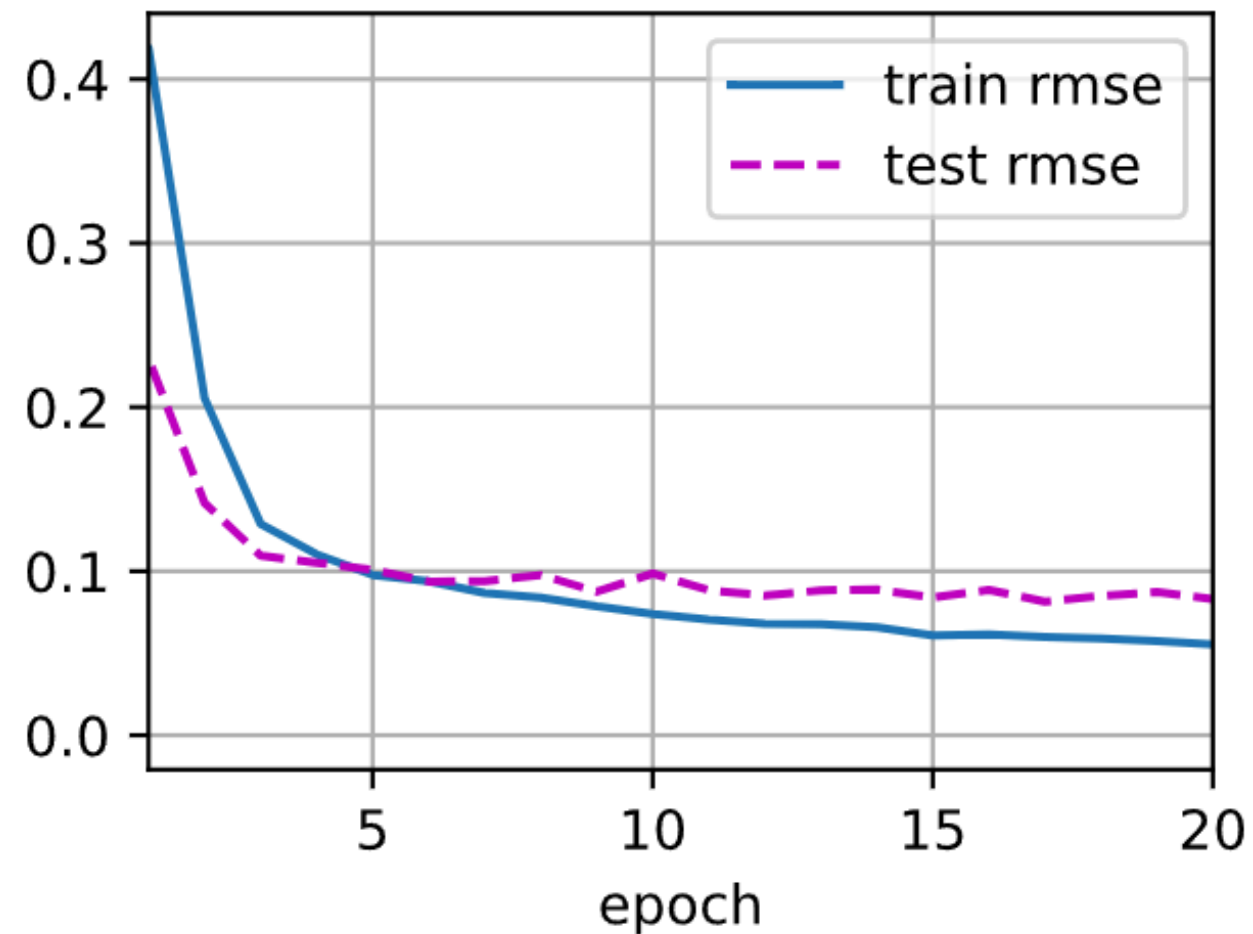
fc: linear layer for regression (1 output channel for regression).

INCEPTION BLOCK





Training



Best test RMSE: 0.08

Optimizer: Adam

It adapts the learning rate for each parameter, improving convergence and stability

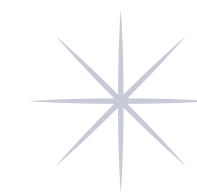
Loss function: MSELoss

MSE Loss measures the average of the squared differences between predicted values and actual values

Learning rate = 0.001

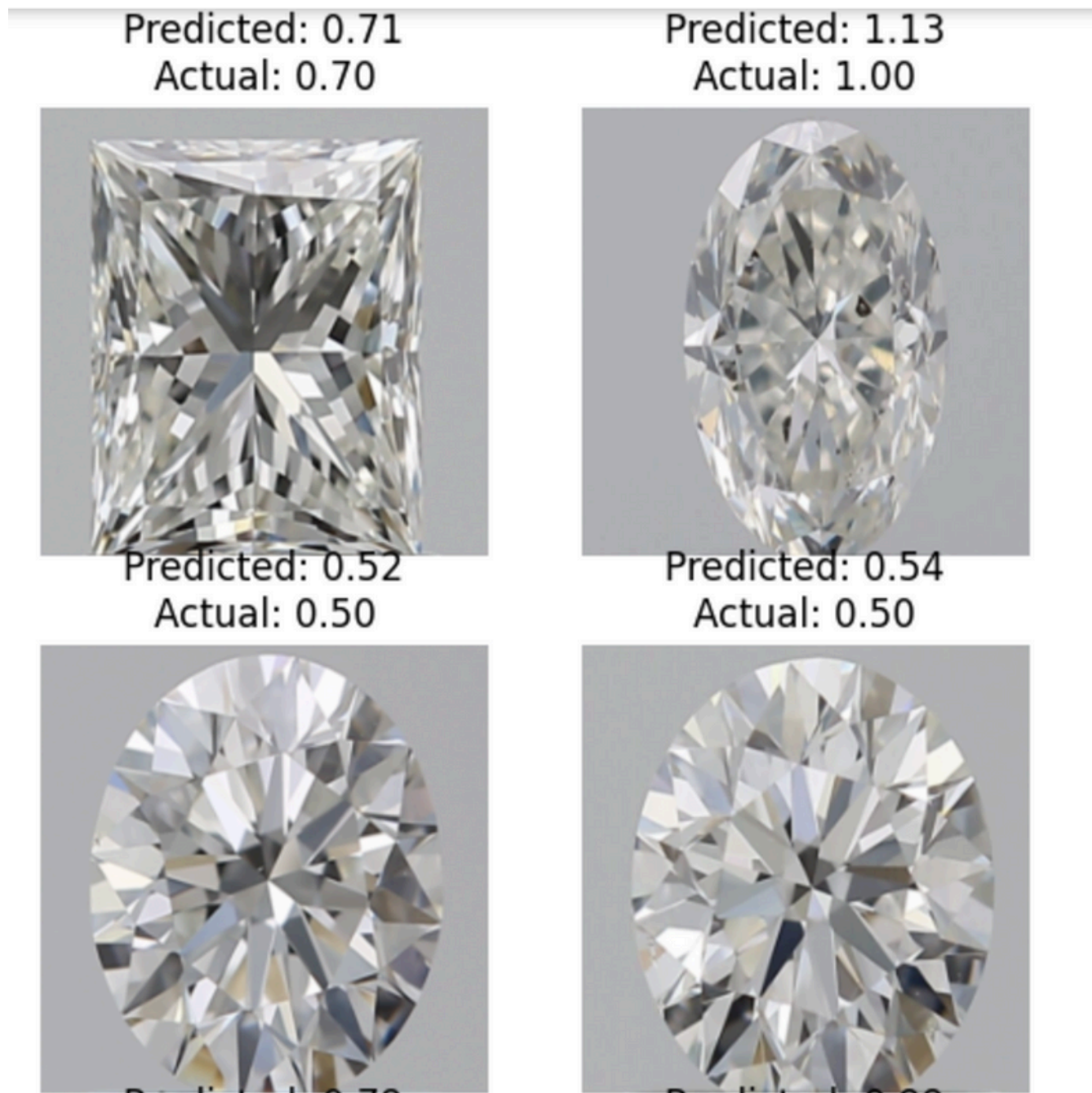
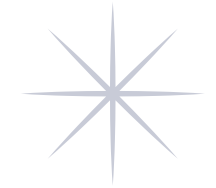
batch size = 64

epochs = 20



Model evaluation on test set

load the model saved during training and use it to make predictions on a sample from the test set



RMSE on test set: 0.08

The final result, with an RMSE of 0.08 on the test set, indicates a good model, despite the challenges posed by the initial dataset. The graphical results show that the model performed well in predicting carat values for diamonds up to 2.2