INTEGRITY • KNOWLEDGE • DILIGENCE
INNOVATIVENESS • PROFESSIONALISM

UNIVERSITY OF ZIMBABWE

**NAME: TAPIWA MHONDIWA**

**REG NUMBER: R229118P**

**MODULE: DEEP LEARNING**

**PROJECT ASSIGNMENT**

**Project Summary: House Price Prediction with Images and Tabular Data**

This mini-project builds a deep learning system to predict residential property prices using both property images and structured features such as building area, land area, number of bedrooms, and location. The goal was to explore whether combining computer vision with tabular data can improve predictive performance for real-estate pricing.

The dataset started with 1,598 property records and 940 images. Records were matched to images using filename and ID, resulting in 460 usable samples. Outliers in the target variable price(USD) were then removed by restricting prices to a reasonable range (10,000–2,000,000 USD), which reduced the dataset to 443 samples and produced a more realistic price distribution for training and evaluation.

Images were processed with standard resizing and normalization, while tabular features were cleaned (missing values handled), encoded (location using label encoding), and standardized. A custom PyTorch Dataset class was implemented to load images and tabular features together and to apply a StandardScaler to the target prices, with train/validation/test splits created at approximately 70/15/15.

The core model is a hybrid architecture that fuses a convolutional neural network (CNN) backbone from torchvision (e.g., EfficientNet, MobileNet-v2, ResNet, VGG, ZFNet) with a small multilayer perceptron for the tabular features. CNN features and tabular features are concatenated and passed through fully connected layers to output a single price prediction. Several training loops were developed, including a final “ultimate” loop using AdamW, learning-rate scheduling, gradient clipping, and a composite loss (MSE, MAE, Huber) for stability.

After cleaning the data, 21 different CNN-based hybrid models were trained and evaluated. On the clean dataset, the best models (such as EfficientNet and MobileNet-v2) achieved test R2*R*2 values in the “GOOD/FAIR” range, with root mean squared error (RMSE) on the order of a few hundred thousand USD and mean absolute error (MAE) around 200k–250k USD, reflecting both model limitations and the high variability of real-estate prices. Some deeper or more complex variants did not converge well, highlighting the impact of small dataset size.

Overall, the project shows that (1) careful data cleaning and outlier handling are critical for stable training, (2) pre-trained CNNs combined with tabular features can learn meaningful price patterns even from a relatively small dataset, and (3) there is still substantial unexplained variance, suggesting that more data, richer location features, and further model tuning would be needed for production-level pricing.