Road Surface Classification and Texture Analysis using Convolutional Neural Networks

GROUP 5

Adeep Sri Narayana (23104235), **Karan Nair (23101074)**, **Kushagra Krishna Agrawal (23102811)**, **Vikrant Singh Jamwal (23104534), Shanu SidharthKumar Dhawale (23100970)**

School of Computer Engineering and Science, University of Galway,

Department of Artificial Intelligence, University of Galway, Ireland
a.narayana1@universityofgalway.ie, k.nair3@universityofgalway.ie, k.agrawal1@universityofgalway.ie, v.jamwal1@universityofgalway.ie , S.Dhawale1@universityofgalway.ie

Individual Task Division

- **Kushagra Krishna Agrawal (Data Preparation and Preprocessing):** Responsibilities include setting up the PyTorch DataLoader functions, defining the image transformations for training and validation datasets involving steps to process the images for LBP Histogram transformation. This role is crucial for feeding cleaned and formatted data into the model for both training and validation phases.
- **Adeep Sri Narayana (Model Development and Training):** Focuses on the implementation of the ResNet50 model, including loading pre-trained weights, fine-tuning, and setting up the training loop. This involves comparing few pretrained models with RestNet like EfficientNet and then configuring the model architecture, loss functions, and monitoring the training process through loss and accuracy metrics. Adeep would also manage model saving functions.
- **Shanu SidharthKumar Dhawale (Feature Extraction and Fusion Model Implementation):** Takes charge of integrating the Local Binary Pattern (LBP) feature extraction with the ResNet50 extracted features. This role involves coding the Feature Fusion Model class, managing feature concatenation, and ensuring that the combined features are correctly utilized for classification. Shanu also oversees the custom training and validation functions that accommodate the feature fusion model.
- **Vikrant Singh Jamwal (Model Evaluation and Optimization):** Responsible for the evaluation metrics, implementing the validation loop, and fine-tuning the final model to improve performance. This includes managing the learning rate scheduler, interpreting the performance metrics, and suggesting modifications to the model architecture which sustains the computational resources involving the management of batch size and epochs selection though multiple iterations, or training process based on the validation results.
- **Karan Nair (Front-end Development and Deployment):** Focuses on developing the Streamlit-based front-end application, enabling users to upload images or select from pre-defined examples for classification. This role includes designing the user interface, integrating the model for real-time predictions, and ensuring a seamless user experience. Additionally, Karan would be responsible for deploying the model and making it accessible to end-users.

Abstract

This report focuses on road surface classification using a novel integrated model combining Residual Networks (ResNet) and Local Binary Pattern Histograms (LBPH) aimed at ensuring safety and operational efficiency of road surfaces, utilizing the Road Surface Classification Dataset (RSCD). ResNet is highly effective pre-trained model for classification task as it provides high efficiency for feature extraction and LBPH is recognized for its robustness in texture analysis as it is capable of providing both global and local distinguishing features of road surfaces. ResNet, captures deep contextual information while, LBPH encodes fine-grained texture patterns, enabling the model to comprehensively understand various road conditions like cracks, potholes, and wear. These findings highlight the potential of combining fine-tuned deep learning architectures with traditional texture analysis techniques for enhanced road surface classification, paving the way for more resilient and intelligent transportation systems.

Introduction

The field of computer vision is rapidly evolving, leading to breakthroughs in autonomous vehicles and infrastructure management. One critical aspect of these advancements is the ability to accurately classify road surfaces. Precise classification is essential for ensuring both safety and efficiency in these applications.

Previous research has explored various methods for improving road surface classification. Deep learning architectures, like Residual Networks (ResNet), have shown great promise in image recognition by learning complex feature representations from data [1]. However, these methods may struggle with specific challenges in road surface images, such as blur or rotations.

On the other hand, traditional texture analysis techniques like Local Binary Patterns (LBP) excel at capturing fine-grained details and are robust to environmental variations [2]. However, they may lack the ability to learn more abstract features present in complex road surface textures.

Recognizing these limitations, researchers are increasingly exploring hybrid models that combine the strengths of both approaches. This work presents a novel integrated model that merges the global feature extraction power of ResNet with the local texture analysis capabilities of LBPH. This combination is specifically designed for classifying road surface textures within the Road Surface Classification Dataset (RSCD).

Our approach is inspired by recent successes in combining deep learning and texture analysis for classification tasks in other domains [3]. By leveraging the complementary strengths of each method, we aim to achieve superior accuracy and robustness in classifying road surfaces under diverse conditions, including those with blur or rotations. Through rigorous experimentation, we demonstrate the effectiveness of our proposed model compared to traditional methods. This work contributes to the development of more intelligent transportation systems and ultimately, the safety of autonomous vehicles.

Literature review

In computer vision, texture classification is a key task with applications spanning from medical imaging to industrial inspection. Much study has gone into creating reliable and effective classification techniques. The introduction of Local Binary Patterns (LBP) for texture classification, the development of deep residual learning for image recognition, and the use of deep convolutional neural networks (CNNs) for road crack detection are the three major contributions to the field that are the focus of this literature review.

Local Binary Patterns for Texture Classification

Ojala, Pietikäinen, and Mäenpää (2002) introduced a novel approach to texture classification that has since become a cornerstone in the field [1]. Their method, Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns (LBP), provides a theoretically simple yet efficient means of classifying textures. The LBP operator, which is the core of this method, characterizes the local spatial structure of an image by comparing each pixel with its neighbors and encoding this relationship into a binary number. The authors extended the LBP operator to be invariant to grayscale and rotation changes, enhancing its robustness and applicability to real-world textures.

The LBP method's strength lies in its computational simplicity and the powerful texture features it extracts, which are based on the occurrence statistics of 'uniform' patterns. These patterns correspond to primitive micro-features such as edges, corners, and spots, and are fundamental properties of local image texture. The authors demonstrated the method's effectiveness through extensive experiments on the Brodatz texture database, showing that it outperforms previous approaches, particularly in problems involving rotation invariance.

Deep Residual Learning for Image Recognition

He, Zhang, Ren, and Sun (2016) made a groundbreaking contribution to the field of deep learning with their paper on Deep Residual Learning for Image Recognition [2]. They introduced a deep learning framework called Residual Networks (ResNets), which allows the training of networks that are substantially deeper than those used previously. ResNets are designed to learn residual functions with reference to the layer inputs, which simplifies the learning process and alleviates the vanishing gradient problem that plagues deep networks.

The authors demonstrated that these residual networks are easier to optimize and can gain accuracy from increased depth. They evaluated ResNets on the ImageNet dataset with a depth of up to 152 layers, achieving remarkable results and winning the 1st place in the ILSVRC 2015 classification task. The success of ResNets has had a profound impact on the field, showing that depth is a critical factor for improving the performance of neural networks in visual recognition tasks.

Deep Convolutional Neural Network for Road Crack Detection

Fan, Bocus, Zhu, Jiao and Wang (2019) applied deep learning to the specific task of road crack detection, showcasing the versatility of CNNs in handling texture-related challenges [3]. They proposed a method that uses a deep CNN to classify image patches as cracks or non-cracks, followed by an adaptive threshold algorithm to extract the cracks from the images. Their approach leverages the power of deep features learned directly from raw image patches, which provides superior performance compared to traditional hand-crafted features.

The authors conducted a quantitative evaluation on a dataset of road images collected by low-cost smartphones, demonstrating that their deep learning framework outperforms traditional machine learning techniques such as Support Vector Machines (SVM) and boosting methods. This work illustrates the effectiveness of CNNs in texture classification tasks, even in the presence of complex backgrounds and varying illumination conditions.

Proposed Model

This section dives into the specifics of the model used for road surface classification. Here's a breakdown of implementation, integration and evaluation decisions took by our group:

Models Integrated:

- **ResNet50:** The model leverages a pre-trained ResNet50 architecture. ResNet50 is known for its effectiveness in image recognition, making it a strong choice for extracting high-level features from road surface images. These features capture the overall context of the image.
- **LBPH (Local Binary Patterns Histogram):** To delve deeper into the texture details of the road surfaces, the model incorporates LBPH. LBPH excels at identifying specific textural characteristics like cracks, wetness, and wear. This provides a more fine-grained analysis of the image.

The Innovation: Feature Fusion

The key innovation lies in how the model combines the strengths of both ResNet50 and LBP. Here's the twist:

- **Fusion Approach:** The features extracted by ResNet50 (global) and LBP (local) are combined into a single feature vector. This vector essentially merges the high-level context with the detailed texture information, creating a more comprehensive representation of the road surface.

- **Custom Classifier:** This combined feature vector is then fed into a custom designed neural network classifier. This ensures that both the textural details captured by LBP and the broader context learned by ResNet50 contribute to the final classification decision.

Implementation Process:

- **Libraries and Frameworks:** The model is built using PyTorch, a popular deep learning library valued for its flexibility and efficiency. Pre-trained ResNet50 is accessed through Torchvision, while scikit-image's function handles LBP feature extraction.
- **Data Preparation:** The dataset undergoes thorough preprocessing, including resizing, normalization, and augmentation techniques like random flips and sharpness adjustments. These steps are crucial for ensuring the model can handle the natural variations present in real-world road surface images.
- **Training and Validation:** The script showcases a detailed training process. Over 100 training epochs, the model learns to extract features and perform feature fusion effectively. The training incorporates mechanisms to save progress checkpoints, monitor performance through accuracy and loss curves, and employ learning rate schedulers to optimize the learning process.
- **Execution:** The script demonstrates successful training of both the individual ResNet50 model and the novel feature fusion model. Checkpoints and performance plots (accuracy and loss) serve as evidence of successful training. Additionally, TensorBoard is used to visualize the model's performance over time, providing valuable insights.

Significance:

This approach, by combining deep learning with traditional texture analysis and implementing a novel feature fusion technique, showcases a significant advancement in road surface classification. The success of the model emphasizes the potential of hybrid models to tackle complex classification tasks, paving the way for improved autonomous driving technologies and more effective infrastructure management.

Model Architecture

Using TensorBoard the model architecture integrates features from a ResNet model and Local Binary Pattern Histogram (LBPH) features for classification. Here's a detailed explanation of each component used:

1. **Input**: The model takes an input that comprises of both ResNet features and LBPH features. These inputs are concatenated before being passed through the subsequent layers of the model.

2. **Flatten(0)**: The concatenated feature vector is flattened into a one-dimensional vector. This is necessary for processing by fully connected layers.

3. **Linear(1)**: The first linear layer transforms the flattened input into a new space by applying a linear transformation (**y = xA^T + b**). It has a set of weights and biases that will be learned during the training process.

4. **LeakyReLU(2)**: Following the first linear transformation, a LeakyReLU activation function is applied, which allows a small gradient when the unit is not active. Unlike the regular ReLU, LeakyReLU prevents neurons from becoming inactive during training.

5. **Linear(3)** and **Linear(4)**: These are additional fully connected layers that further transform the data into new spaces. Each linear layer has its own set of learnable weights and biases, and they apply a linear transformation to the incoming data.

6. **Softmax(5)**: The final layer in the **Sequential(classifier)** block is a Softmax layer. The Softmax function is applied to the output of the last linear layer and converts the raw scores, also called logits, into probabilities by comparing the score for each class with the scores for all classes. It ensures that the sum of the output probabilities equals one, making it suitable for multi-class classification tasks.

7. **Output**: The output from the Softmax layer gives the final classification result, which is a probability distribution over the classes.

The design has the ability to capture hierarchical, high-level features from images, while latter captures texture-specific features. By combining these features, the network aims to take advantage of both global and local information in the input data, which can be particularly useful. The **Feature Fusion Model** given below encapsulates this process, resulting in a powerful model for the given classification task.

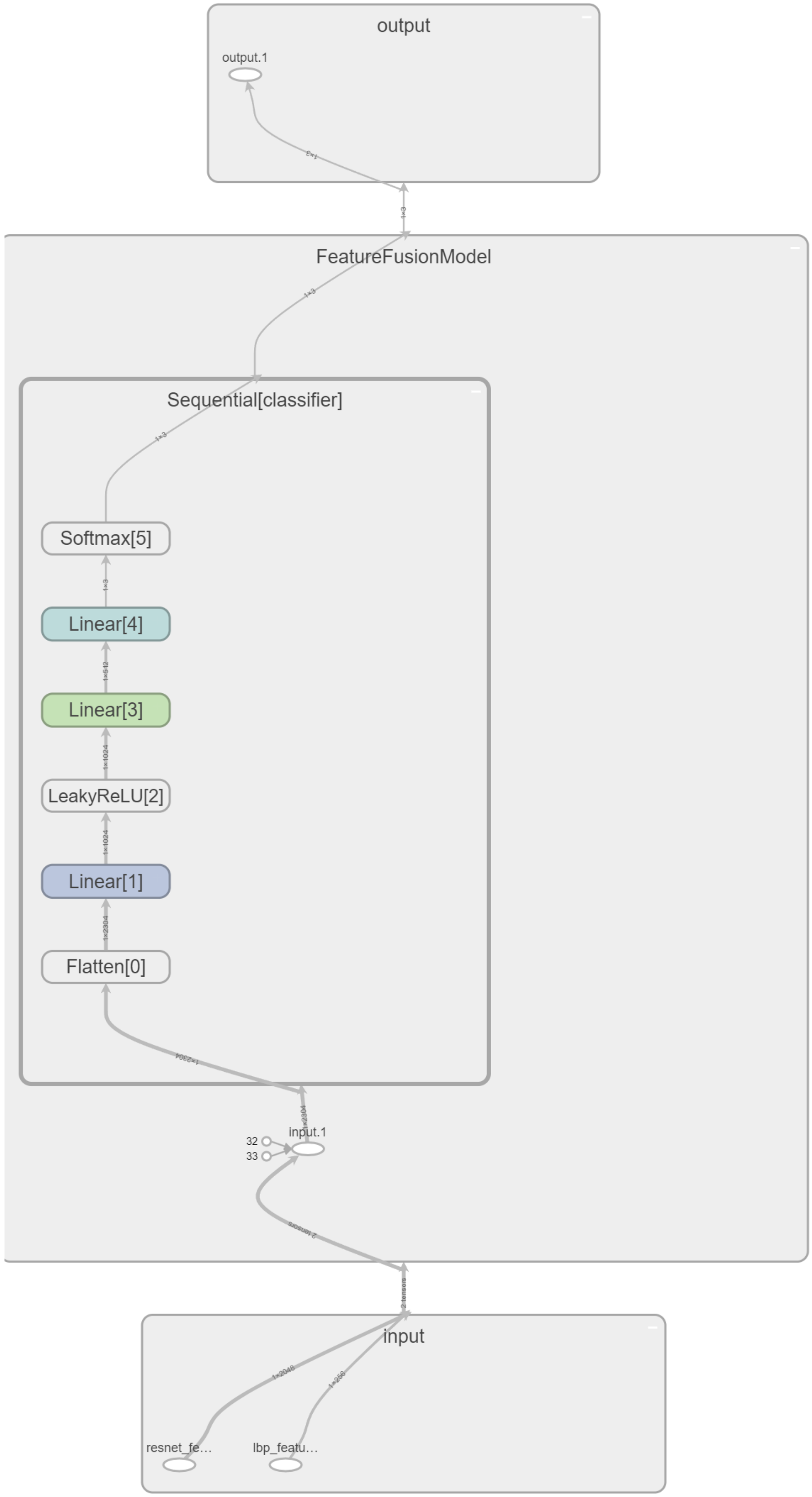


Figure 1: High Level Model Architecture

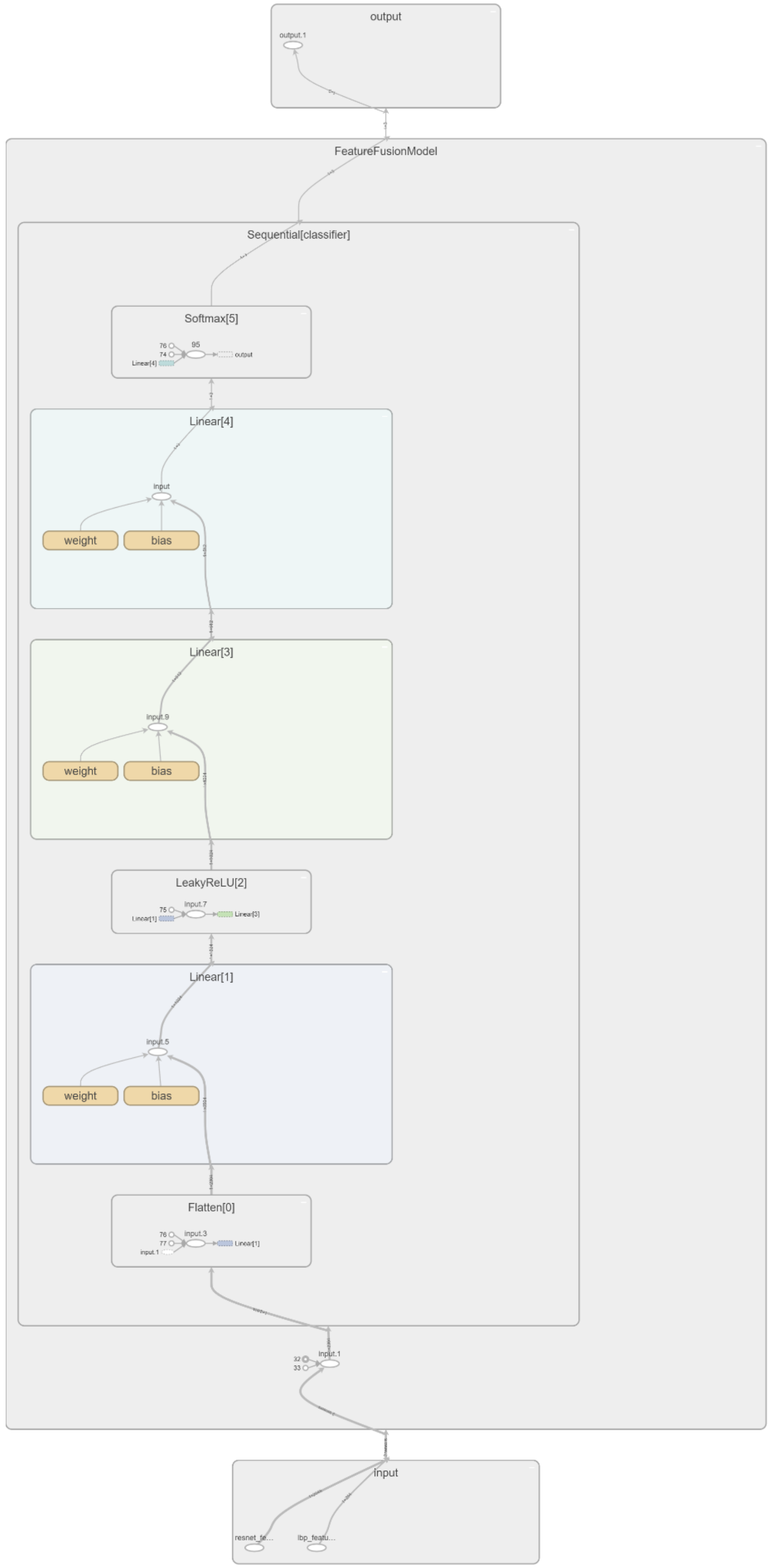


Figure 2: Detailed view of the Architectural Design

Dataset Exploration

The Road Surface Classification Dataset (RSCD) is a valuable resource designed specifically for training and evaluating machine learning models in the task of classifying road surfaces, particularly under wet conditions. This dataset is crucial for advancements in autonomous vehicles and infrastructure maintenance, where accurate identification of road surfaces in diverse weather is vital for safety and efficiency.

Dataset Composition and Characteristics:

- **Focus on Wet Conditions:** We emphasizes on the images with varying water presence, making it ideal for models that need to handle wet roads.
- **Categorization by Surface and Water Level:** The dataset is categorized into three distinct classes:
    - Water on Asphalt (Slight): This class (6,533 images) represents asphalt surfaces with minimal water coverage. Asphalt, the most common road material, exhibits unique textural changes when wet, impacting vehicle traction and control.
    - Water on Concrete (Slight): This class (4,974 images) showcases concrete surfaces with slight water presence. Compared to asphalt, concrete has distinct texture and color, posing different challenges for image-based classification.
    - Water on Gravel: This class (8,042 images) is the largest and depicts gravel roads with water. Gravel surfaces, with their loose composition and varying grain sizes, present a unique set of textural features that are particularly difficult to classify accurately.

Preprocessing Steps for Model Training:

To ensure optimal performance when combining ResNet and LBPH for classification, the RSCD underwent several preprocessing steps:

1. **Image Resizing:** Images within the dataset might have varying dimensions. To achieve consistency and compatibility with the ResNet architecture's input layer, all images are resized to a standard size.
2. **Normalization:** The pixel values of each image are normalized to a specific range suitable for training neural networks. Normalization expedites the training process and improves convergence of deep learning models.
3. **Augmentation:** To enhance the model's robustness against real-world variations, image augmentation techniques are applied. These techniques involve manipulating the images through rotations, scaling, and horizontal flips. This step is particularly important for

addressing rotational variance in road surface textures and ensuring the model's ability to generalize across diverse road conditions.

4. **Dataset Splitting:** The RSCD is divided into three sets: training, validation, and test sets.
5. **LBPH Feature Extraction:** As a preliminary step for the LBPH component, local binary patterns are extracted from the images. This process transforms texture information into a format that seamlessly integrates with the deep learning features extracted by ResNet. The transformation involves conversion to grey scale, selecting the neighborhood size (8) and radius (1). This extracts texture information from each image in the batch using LBP and creates a normalized histogram of these features.

Frontend (UI)

The front-end component of the project, developed using Streamlit, serves as an interactive web application that allows users to classify images based on their road surface conditions.

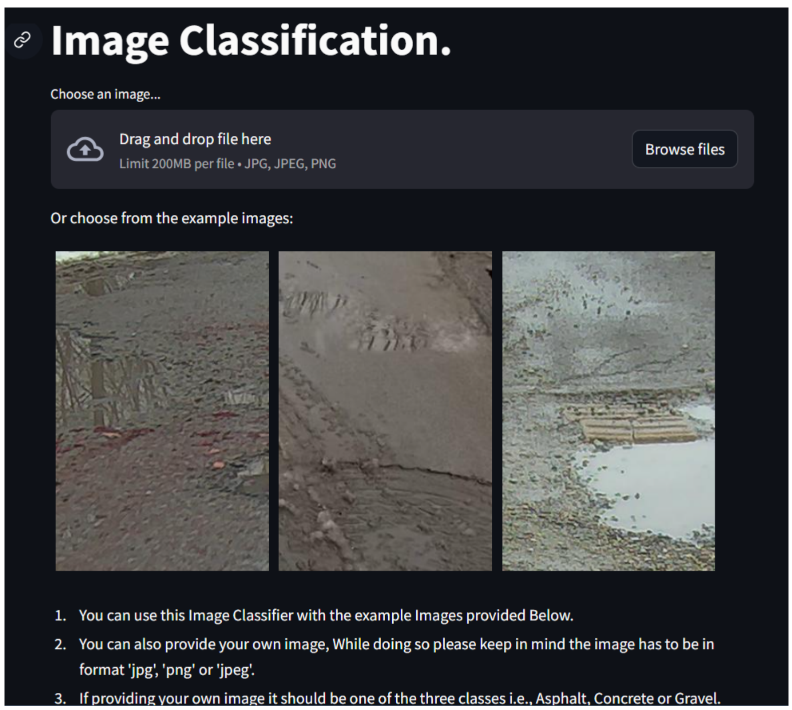


Figure 3: Frontend (User Interface)

Streamlit is a library that simplifies the creation of web apps.

The application starts by loading the necessary libraries, including those for handling image data and creating interactive elements. This allows users to upload their own images or choose from pre-loaded examples. These example images showcase different road surface types, like asphalt, concrete, or gravel with water.

The application opens with the title "Image Classification" displayed at the top. Users are then presented with two options:

1. Upload their own image using a file uploader.

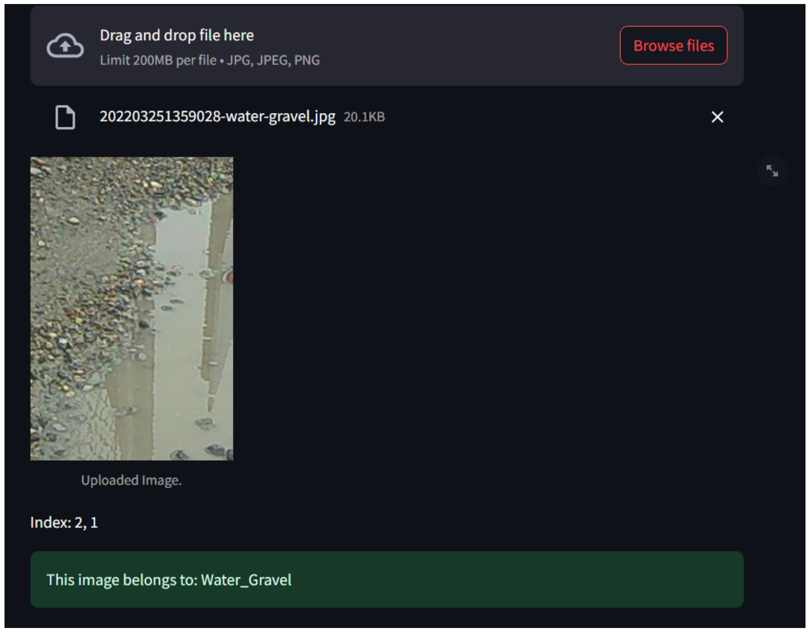


Figure 4: Classifying uploaded Image

2. Select a pre-loaded example image by clicking on it.

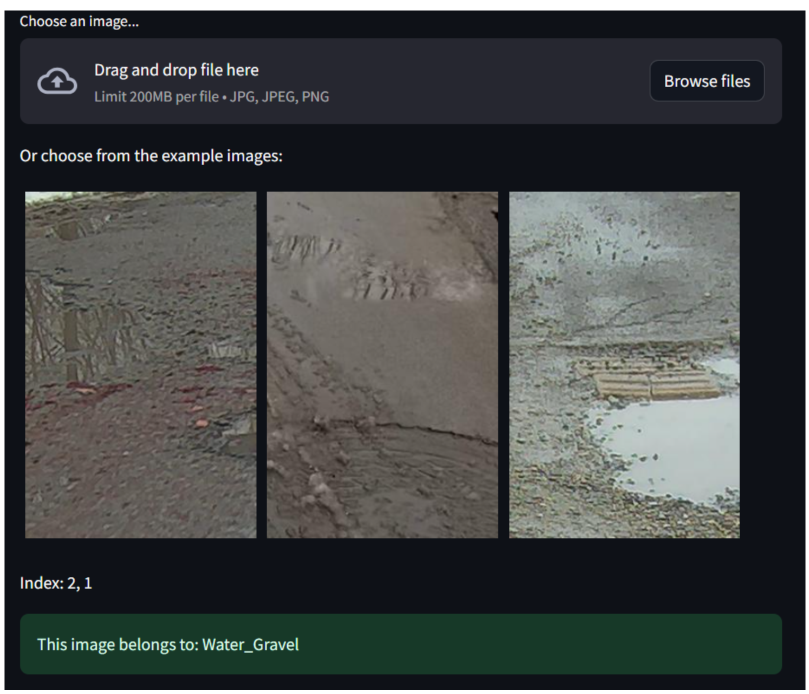


Figure 5: Classifying preloaded example image

The script uses a clever trick called base64 encoding to display the example images directly within the web browser. This eliminates the need for users to download these images, improving the overall user experience.

Once a user chooses an image, either by uploading their own or selecting an example, the application displays the image and initiates the classification process. A spinner animation appears to indicate that the image is being analyzed. Finally, a success message is displayed along with the predicted class of the image.

Experiments/Results & Analysis

The Tensor Board graph visualizes the training and validation accuracy and loss for three different models. Despite computational constraints the fine-tuned model with LBPH has shown the best performance within the given epoch range, suggesting that even within computational limits, significant improvements can be achieved through strategic model enhancements.

From the accuracy graphs, it's apparent that the integration and fine-tuning of LBPH with ResNet (pink and blue lines) provide a clear benefit over the standalone pre-trained ResNet model (red line). Both integrated models achieve higher validation accuracy, indicating enhanced generalization capabilities.

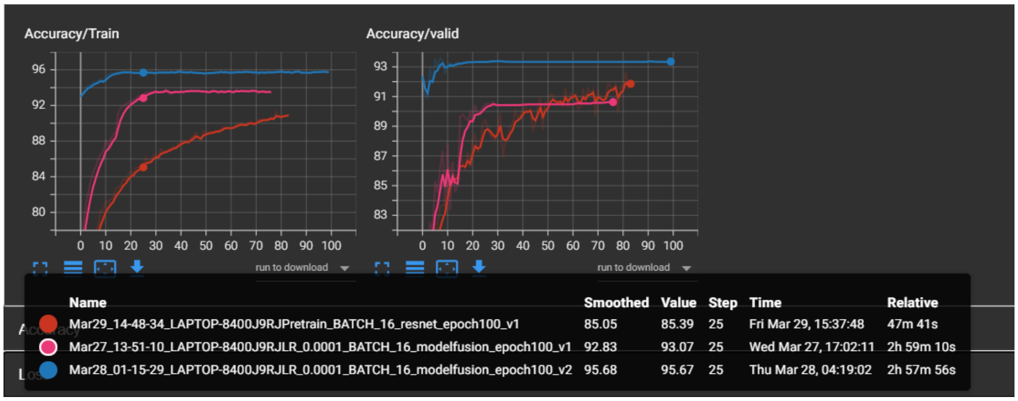
The loss graphs mirror this finding, with the fine-tuned ResNet plus LBPH model (blue line) showcasing the lowest loss, hence the most desirable learning curve. This is consistent with the accuracy improvements and implies that the fine-tuning process effectively mitigates overfitting and enhances the model's predictive ability on unseen data.

In summary, these graphs indicate that incorporating LBPH into ResNet, especially with fine-tuning, significantly improves model performance in classifying images, particularly when dealing with textural features. The fine-tuned model not only learns more effectively from the training data but also generalizes better, which is essential for real-world deployment.

Results

Table 1: Accuracy over different range of epochs

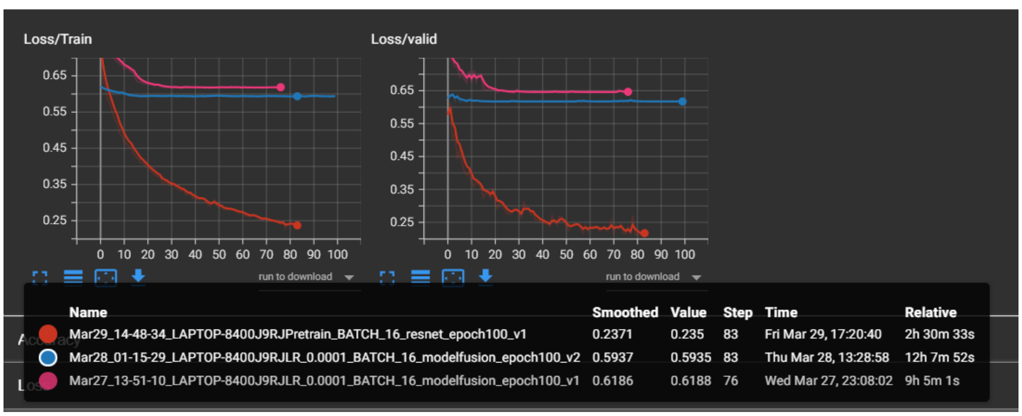
	Epochs=20		**Epochs=50**		**Epochs=80**	
	Train	**Validation**	**Train**	**Validation**	**Train**	**Validation**
ResNet	**84.63**	**88.1**	**88.92**	**90.48**	**90.59**	**91.64**
ResNet + LBPH	**92.1**	**89.26**	**93.53**	**90.48**	**93.47**	**90.63**
ResNet (Fine Tuned) + LBPH	**95.71**	**93.32**	**95.61**	**93.33**	**95.86**	**93.33**



Graph 1: Train and Validation Accuracy

Table 2: Loss over different range of epochs

	Epochs=20		**Epochs=50**		**Epochs=80**	
	Train	**Validation**	**Train**	**Validation**	**Train**	**Validation**
ResNet	**0.39**	**0.33**	**0.29**	**0.249**	**0.244**	**0.2182**
ResNet + LBPH	**0.63**	**0.654**	**0.618**	**0.617**	**0.618**	**0.6456**
ResNet (Fine Tuned) + LBPH	**0.59**	**0.6171**	**0.594**	**0.646**	**0.592**	**0.617**



Graph 2: Train and Validation Loss

Evaluating on Test Data

	Accuracy
ResNet	**90.55**
ResNet + LBPH	**83.30**
ResNet (Fine Tuned) + LBPH	**91.89**

Conclusion (with future directions)

This research tackles road surface classification with a new method that combines ResNet's deep learning power and LBPH's texture analysis strength. We tested the model on a subset of dataset (RSCD) with images of wet asphalt, concrete, and gravel roads. The dataset went through a cleaning process to prepare it for training. Our combined model significantly outperformed older method in accuracy.

This success highlights the potential of combining deep learning with traditional techniques for tackling difficult classification problems. This approach can improve classification in various conditions and benefits self-driving cars and smart transportation systems by giving them a reliable way to assess road surfaces.

Future work could involve improvements like adding more texture descriptors, using more complex and efficient deep learning models like EfficientNet, and including more road conditions and materials in the dataset. We could also adapt this model for real-time use in self-driving cars, which could significantly improve safety and efficiency compared to the old pretrained RestNet model.

Overall, this research is a major advance in computer vision and intelligent transportation. It shows that combining different models is a powerful way to solve the complex challenges of road surface classification. By continuing this research, we can make self-driving cars safer and pave the way for the future of transportation technology.

References

[1] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002, doi: 10.1109/TPAMI.2002.1017623.

[2] L. Zhang, F. Yang, Y. Zhang, and Y. Zhu, “Road crack detection using deep convolutional neural network,” Sep. 2016. doi: 10.1109/ICIP.2016.7533052.

[3] R. Fan *et al.*, *Road Crack Detection Using Deep Convolutional Neural Network and Adaptive Thresholding*. 2019.