FixCyprus: Automated Classification of Crowdsourced Reports Using Machine Learning

1st Leslie Miller
SenSIP Center, ECEE
Arizona State University
Tempe, USA
lmille42@asu.edu

2nd Christos Laoudias
KIOS
University of Cyprus
Aglantzia, Cyrpus
laoudias.christos@ucy.ac.cy

3nd Christos Kyrkou
KIOS
University of Cyprus
Aglantzia, Cyrpus
kyrkou.christos@ucy.ac.cy

4rd Andreas Spanias
SenSIP Center, ECEE
Arizona State University
Tempe, USA
spanias@asu.edu

***Abstract*—In this research, we utilize the FixCyprus dataset, a cost-effective crowdsourcing service that allows road transportation authorities in Cyprus to gather information about the road network and surrounding infrastructure. Users can upload reports on issues such as damage, traffic signs, and road blockages through the FixCyprus app, among others. Each report includes an image and various information including report category, infrastructure category, location, priority level, and more. In this study, we begin with a detailed analysis of the dataset, including image information and label frequency statistics. Next, we utilize various machine learning algorithms to classify the data based on the Report Category. More specifically, we design and deploy a ResNet-50 model for feature extraction, incorporating the extracted features into classifiers such as SVM, KNN, and Decision Trees. The performance of the model is evaluated using F1 scores and classification accuracy.**

***Index Terms*—Machine Learning, Crowdsourcing, Classification**

I. Introduction

In this report, we detail the use of the FixCyprus dataset to classify data based on reports from citizens of Cyprus. This app was introduced in 2022 to provide citizens with a smarter and safer road network. By the end of March 2023, the FixCyprus app had over 6,000 users, and around 16% of all registered users had reported at least one issue [1]. In this study, we examine the data submitted by the users. We provide an analysis of the submitted data and attempt to develop a machine-learning model to classify the images according to the Report Category.

II. Literature Review

Paper [1] describes the creation of the FixCyprus dataset. The Ministry of Transport, Communications and Works (MTCW) in Cyprus is closely collaborating with the KIOS Research and Innovation Center of Excellence (KIOS CoE) at the University of Cyprus to provide a safer and smarter road network to the citizens of Cyprus. In this paper, the authors describe the functionality of the FixCyprus app. The FixCyprus app is a mobile app that can be downloaded from the Apple/Google store. Once logged in, users can either create a new incident report or browse previous reports in the history view. To create a new report, users take a geotagged photo of the issue, select the incident category (e.g., damage, obstacle, vandalism), and identify the relevant infrastructure (e.g., road, pavement, cycle path). Users then confirm or manually adjust the incident location within a fixed range. Finally, they can add comments before submitting the report, which PWD officers will review. The overall process is illustrated in Figure 1 [1]. In addition to the above, the authors of [1] also demonstrate the two modes in the app: Interactive and Trace Driven mode. They also describe applications where ML could be useful, such as detecting and flagging irrelevant images uploaded by users.

The rest of the paper is organized as follows. First, we include a literature review. In Section III, we describe the dataset analysis performed, in Section IV we our ML algorithm, in Section V we report our results, and in Section VI we provide concluding remarks.

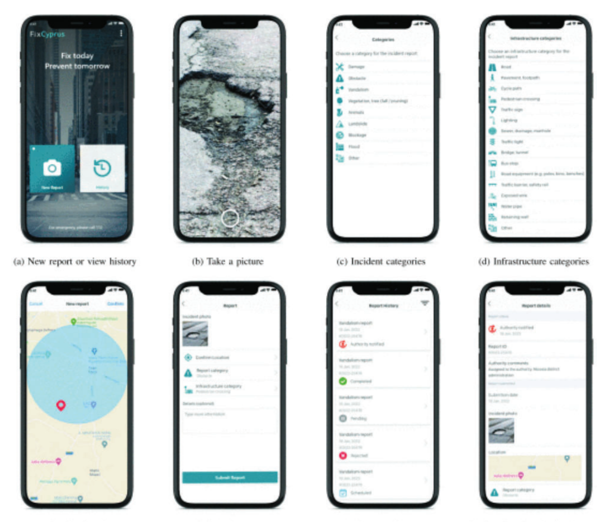

Fig. 1: User Interface for the FixCyprus Application [1].

In this study, we are focused on classifying images based on the Report and Infrastructure Categories. Various ML algorithms are useful when performing image classification. This includes convolutional neural networks, feature extractors combined with classifiers like SVMs and RNNs, and other techniques. Paper [2] uses CNNs to extract features from their dataset, which includes both paved and unpaved roads, for a Road Condition Inspection System and employs SVM and Random Forest for classification. They use varying image sizes ranging from 50x50 to 200x200. The best-performing image size and classifier was 200x200 and Random Forest. They reported a classification accuracy of 96.2%. Their worst-performing model was when they used an image size of 50x50 and a Random Forest classifier. In this methodology, they achieved a classification accuracy of 51.8%.

III. Dataset Analysis

For the purpose of this study, we first begin with a statistical analysis of our image dataset. Each report submitted to the FixCyprus app includes an image and various information including report category, infrastructure category, location, priority level, and more. We analyze various types of data, including the frequency of labels across different categories and pixel information. In this specific study, we focused on the report category for scene classification. This category comprises nine labels: Animals, Blockage, Damage, Flood, Landslide, Obstacle, Vandalism, Vegetation, and Tree. The entire dataset includes 7,724 images and data entries in the CSV file. Figure 2 illustrates the distribution of the labels in the Report Category, with 'Damage' being the most frequently reported and 'Landslide' the least. Figure 3 shows an example image from each of the nine classes in the Report Category.

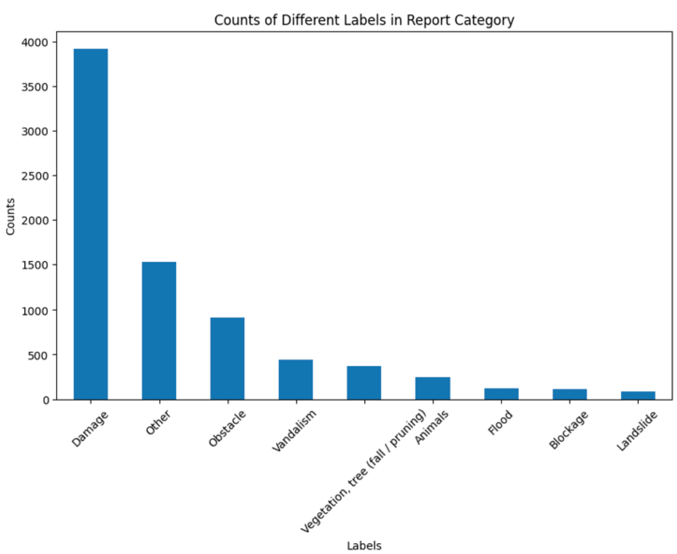

Fig. 2: Counts of Different Labels in the Report Category.

Fig. 3: Representation of Dataset in the Report Category.

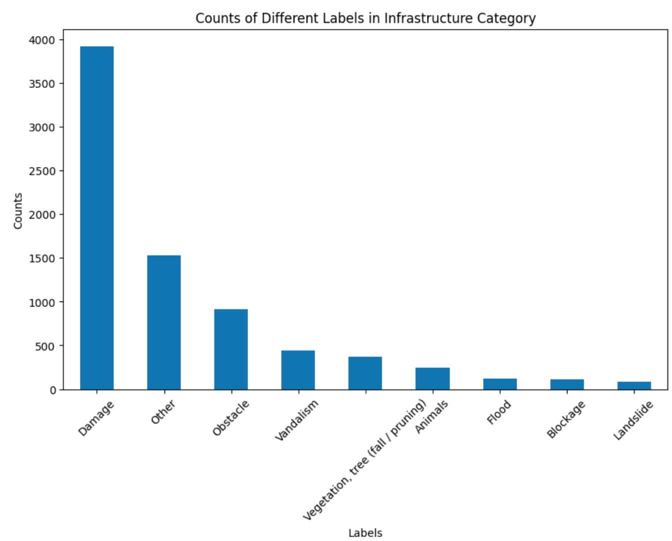

Fig. 4: Counts of Different Labels in the Infrastructure Category.

Figure 4 shows the distribution of labels in the Infrastructure Category. We also examined the range of pixel dimensions in the dataset. The smallest image submitted to the FixCyprus app was 400x533 pixels, while the largest was 4200x5600 pixels. The mean pixel dimensions of all submitted images were 2190x2190, and the median dimensions were 2424x2424. Figure 5 shows the distribution of image size for the entire dataset.

IV. Machine Learning Algorithm

Next, we looked at various machine-learning techniques to classify our data [3,4]. We first began this study looking at using various convolutional neural networks to perform classification. This algorithm did not perform well on the data. We shifted to using the ResNet-50 as a feature extractor and used various classifiers for classification. More specifically, we use a Support Vector Machine (SVM), K-nearest neighbors (KNN), and Decision Trees for classification. The overall architecture can be seen in Figure 7, where we have a ResNet-50 feature extractor and an SVM, KNN, or Decision Tree classifier. The resulting image is assigned a class label that corresponds to the Report Category. We obtain F1 scores ranging from 0.36 to 0.94 based on the model and classes chosen. Our classification results can be found in Section V.

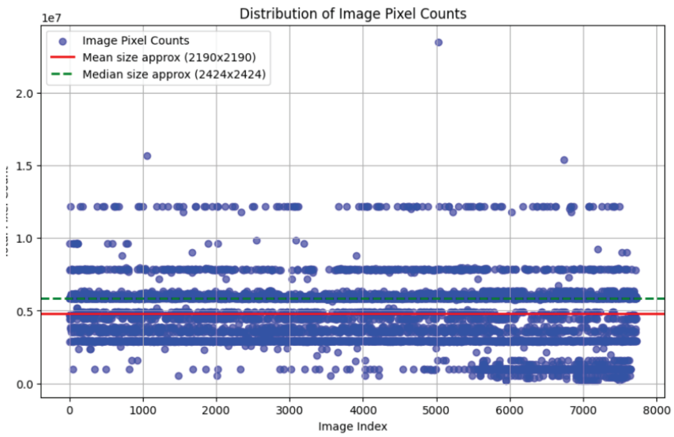

Fig. 5: Distribution of Image Size Across the Entire FixCyprus Dataset

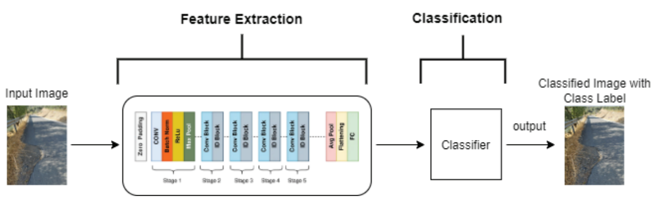

Fig. 6: Architecture for Classification of the Report Category

V. Discussion of Results

Section A presents our findings, while Section B outlines the challenges encountered during the machine learning process and provides recommendations for data improvement to enhance the training of our ML model.

A. Results

Table 1, Table 2, and Table 3 detail the results for classification using feature extraction + an SVM, a KNN, and feature extraction + a KNN. We show eleven different binary combinations of labels in Table 1 and report the Weighted F1 scores along with classification accuracy. Out of the eleven combinations shown, the best performing was Damage with Blockage. This achieved a Weighted Average F1 score of 0.94 and an accuracy of 0.94. The worst performing in this subset was Landslide with Vandalism. This combination's Weighted Average F1 score was 0.69 and classification accuracy was 0.75. Next, we performed multi-class classification using a KNN and feature extraction + KNN. We used three different combinations of four classes; the results can be seen in Table 2 and Table 3. Using a feature extractor coupled with a KNN performed better than just a KNN alone for all three combinations. The best-performing combination of classes was Animals, Damage, Obstacle, and Vegetation with a Weighted Average F1 score of 0.59, and accuracy of 0.63. Table 4 presents the results of using a feature extractor with Decision Trees as the classifier, while Table 5 displays the outcomes when using Random Forests with 15 trees. Overall, the best-performing methodologies were the ResNet50 feature extractor combined with a KNN.

TABLE I: F1 Score and Accuracy Results for Different Combinations of Classes using an SVM

Classes	Weighted Average F1 Score	Accuracy
Landslide, Vandalism	0.69	0.75
Damage, Obstacle	0.70	0.69
Obstacle, Vandalism	0.73	0.73
Flood, Vandalism	0.75	0.78
Blockage, Vegetation	0.77	0.77
Landslide, Obstacle	0.80	0.76
Blockage, Vandalism	0.81	0.83
Blockage, Obstacle	0.84	0.85
Animals, Vandalism	0.85	0.85
Animals, Vegetation	0.88	0.88
Damage, Blockage	0.94	0.94

TABLE II: F1 Score and Accuracy Results for Different Combinations of Classes using a KNN

Classes	Weighted Average F1 Score	Accuracy
Animals, Blockage, Obstacle, Vegetation	0.36	0.32
Damage, Obstacle, Vandalism, Vegetation	0.57	0.68
Animals, Damage, Obstacle, Vegetation	0.59	0.63

TABLE III: F1 Score and Accuracy Results for Different Combinations of Classes using a ResNet-50 Feature Extractor and KNN Classifier

Classes	Weighted Average F1 Score	Accuracy
Animals, Blockage, Obstacle, Vegetation	0.57	0.58
Damage, Obstacle, Vandalism, Vegetation	0.60	0.67
Animals, Damage, Obstacle, Vegetation	0.65	0.69

TABLE IV: F1 Score and Accuracy Results for Different Combinations of Classes using a ResNet-50 Feature Extractor and Decision Tree Classifier

Classes	Weighted Average F1 Score	Accuracy
Animals, Blockage, Obstacle, Vegetation	0.47	0.49
Damage, Obstacle, Vandalism, Vegetation	0.56	0.56
Animals, Damage, Obstacle, Vegetation	0.58	0.59

In addition to the above, to address the imbalance in the 'Damage' category, we developed a feature extractor with the SVM classifier to distinguish between 'Damage' and all other categories. The F1 score and accuracy results for this approach are presented in Table 6, while Figure 7 displays the confusion matrix. We report a Weighted Average F1 score of 0.61 and classification accuracy of 0.61.

TABLE V: F1 Score and Accuracy Results for Different Combinations of Classes using a ResNet-50 Feature Extractor and Random Forest Classifier

Classes	**Weighted Average F1 Score**	**Accuracy**
Animals, Blockage, Obstacle, Vegetation	0.60	0.68
Damage, Obstacle, Vandalism, Vegetation	0.56	0.66
Animals, Damage, Obstacle, Vegetation	0.63	0.71

TABLE VI: F1 Score and Accuracy Results for Damage and Everything Else using a ResNet-50 Feature Extractor and SVM

Classes	**Weighted Average F1 Score**	**Accuracy**
Damage, Everything Else	0.62	0.68

B. Challenges and Recommendations

There were many challenges encountered in this study, which are listed below. In addition to the list of challenges, we also give a few recommendations to enhance the image data, making it more suitable for ML-related tasks.

- Unbalanced dataset (65% of data in the Damage Category)
- Small dataset
- Ambiguity in Difference of Labels (Blockage vs. Obstacle)
- Variation in Image Size
- Long training times when using CNNs
- Not applicable images (random images that do not fall into one of the report categories)

We faced significant class imbalance in our small dataset, along with ambiguity in the labels. For instance, ’blockage,’ ’obstacle,’ and ’animal’ could be interpreted as the same category depending on the user. Additionally, there was considerable variation in image sizes. To meet the requirements of our ML model, we resized all images to 224x224, which resulted in some loss of information. Furthermore, numerous images submitted did not fit into any of the predefined report categories.

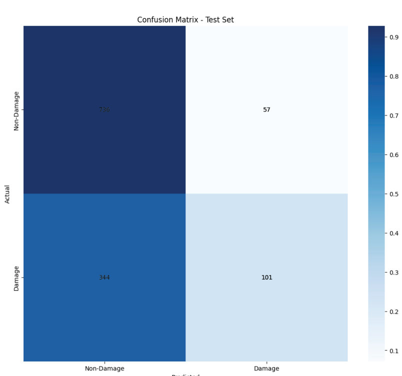

Fig. 7: Confusion Matrix Damage vs. Non-Damage using an SVM.

Recommendations for improving the data include reducing the number of labels to minimize ambiguity. The app should automatically resize all images to a uniform size. Additionally, users should be prompted to ’zoom in’ on the infrastructure-related issue. By using both the zoomed-out and zoomed-in images, we can enhance the training of our model.

VI. Conclusion

In conclusion, we conducted a statistical analysis of the dataset and developed several machine-learning algorithms for classification. We explored a ResNet feature extractor combined with classifiers such as SVM, KNN, and Decision Tree, reporting F1 scores ranging from 0.36 to 0.94. The performance was notably better when using the ResNet feature extractor. Additionally, we performed binary classification of ’Damage’ versus all other categories using an SVM, achieving an F1 score and accuracy of 0.61. We discussed the results, identified challenges, and provided recommendations to improve classification accuracy. In addition to the above, we presented a paper titled ”Image Fusion and Quantum Machine Learning for Remote Sensing Applications” at the IISA conference in Crete, Greece and had a paper accepted to the IEEE MLSP conference at Imperial College London. Future steps include developing a Quantum SVM for classification and continuing to fine-tune the current ML model.

Acknowledgment

Support for this project was provided by the Quantum Collaborative, the SenSIP center, and the NSF IRES Grant Award #1854273.

References

[1] G. Christou, A. Georgiou, C. Laoudias, A. Savva and C. G. Panayiotou, ”FixCyprus: Crowdsourcing Smartphone Imagery Data For Managing Road Safety Hazards,” 2023 24th IEEE International Conference on Mobile Data Management (MDM), Singapore, Singapore, 2023, pp. 188-191.

[2] V. Pereira, S. Tamura, S. Hayamizu and H. Fukai, ”Classification of Paved and Unpaved Road Image Using Convolutional Neural Network for Road Condition Inspection System,” 2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA), Krabi, Thailand, 2018, pp. 165-169.

[3] Sartika, Z. Zainuddin and A. A. Ilham, ”Detection and Classification of Road Damage Using Camera with GLCM and SVM,” 2023 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), BALI, Indonesia, 2023, pp. 372-376, doi: 10.1109/IAICT59002.2023.10205539.

[4] Winarto and D. H. Widyantoro, ”Road and Roadside Detection using Machine Learning Approaches,” 2019 International Conference on Data and Software Engineering (ICoDSE), Pontianak, Indonesia, 2019, pp. 1-5, doi: 10.1109/ICoDSE48700.2019.9092696.