An enhanced approach for Plant Leaf Disease Detection

Akash Bhakat [a], Naveen Nandakumar [b] and Rajkumar S [c]

[a] Vellore Institute of Technology, India
[b] Vellore Institute of Technology, India
[c] Vellore Institute of Technology, India

Abstract

In the modern world with ever rising population, the demand and stress on the agricultural produce to meet the ever growing demand is on record-level high. With the innovations like Smart Agriculture, Natural Fertilizers and Genetically Modified Plants, there still are various areas to focus upon.One of the main areas to focus that is majorly gone unnoticed is the diseases in the plants. In many developing countries with less access to human healthcare there is little to no scope to consider for plant health and disease detection. Due to this a large amount of produce is lost due to diseases in plants. Furthermore most of these diseases can spread from one plant to another starting a domino effect destroying the entire fields.Though some work on this area, the accuracy achieved by the systems can still be increased and systems be made easier to incorporate, use. This paper aims to solve the problem of Detection of Plant Disease by analyzing the image of plant leaves. It aims to increase the existing accuracy of the various existing works with the proposed approach using Transfer Learning and identifying disease in a broader number of leaves than the existing works.

Keywords

Leaf Disease, CNN, transfer learning, AI, farming, plant diseases

1. Introduction

In this modern world, India still depends a lot on Agriculture. Agriculture provides 17% of the GDP and provides employment to more than 60% of the Population. Also this sector is responsible for feeding and providing with various Raw Products for Agro-Based Industries to satisfy the needs of the 1.36 Billion People in India.

With the ever growing demand and the land for the supply being limited and even depleting, it is very much required to make the best use of the resources like land, water. For improving the quantity and quality of the products, many innovative technologies like the use of Genetically Modified Plants produce more Produce per Plant. But still there are huge losses endured due to diseases in plants.

Roughly the losses in agricultural production due to pathogens, diseases, weeds 20% to 40% of the global production. Pathogens and pests are causing 10 percent to 28 percent losses in wheat, 25 percent to 41 percent losses in rice, 20 percent to 41 percent losses in maize, 8 percent to 21 percent losses in potato, and 11 percent to 32 percent losses in soybeans on a global scale, according to a study published in the journal Nature, Ecology & Evolution. In this paper we acknowledge and focus on the disease of the Plants. We believe that the diseases of the plants if detected earlier can help to contain the spread of the disease and also help to produce more.

Algorithms, Computing and Mathematics Conference, August 19 – 20, 2021, Chennai, India.
EMAIL: rajkumars@vit.ac.in (Rajkumar S)
ORCID: 0000-0001-5701-9325 (Rajkumar S)

In this paper, we aim to take up this issue of losses due to diseases in plants and propose a method of solving it. So, in this paper, we aim to study the existing models and propose a model which will help in identifying disease with good accuracy in the leaves of Tomato, Apple, Blueberry, Cherry, Grapes, Corn, Orange, Peach, Raspberry, Soya bean, Squash, Strawberry.

The proposed work focuses on 26 different categories of diseases that occur in the mentioned 12 types of plants.

2. Literature Review of Existing Methods

Table 1: Review of various classification methods of leaves

Author	Method used	Drawbacks
[1] Ghaiwat et al (2014)	ANN, SVM, fuzzy logic. etc	When training data is not linearly separable, it is difficult to grasp the structure of the algorithm and identify appropriate parameters in neural networks.
[3] Badnakhe et al (2011)	KNN with neural network for detecting the diseases on leaves automatically	Crop diseases may be classified using artificial neural networks, fuzzy logic, and other soft computing techniques.
[4] Arivazhagan, S. et al(2013)	Color co-occurrence method with SVM	The training samples can be increased, along with shape and color characteristics, as well as the best features, can be used as illness identification input conditions.
[7] Naikwadi et al (2013)	The spatial gray-level dependency matrices were used to build the color co-occurrence texture analysis technique.	Results could be easily improved with a larger database and advanced feature of color extraction
[11] Mondal et al(2015)	Color Co-Occurrence technique, k-means clustering, Bayesian classifier	Used 12 features for classification but the overall accuracy was only 87%
[12] Padol et al (2016)	Color co-occurrence technique, K-means clustering algorithm using SVM.	The accuracy could have been improved by using fusion classification techniques. The current accuracy remains to be 88,9%.
[13] Reza et al (2016)	Color co-occurrence methods, Multi SVM classifier.	Uses a multi-SVM classifier which gives an accuracy of 86%
[14] Tejonidhi et al(2016)	Bhattacharya’s distance method	It recognizes paddy's burning and blast diseases. In addition, this technique may be used to detect a variety of illnesses in different leaves. This might aid farmers in identifying the illness in the leaf in a practical and precise manner in a short period of time.
[17] Pawar et al (2016)	GLCM(Gray level co-occurrence method), ANN	Using extra texture characteristics can improve classification accuracy. The current accuracy of the study remains to be 80.45%.
[18] Narmadha et al (2017)	Color co-occurrence technique, ANN, FUZZY	Uses KNN and the accuracy achieved was 94.7%.

	classification, SVM, KNN	
[19] Tripathi et al (2016)	K-means, GLCM, ANN, SURF, CCM, SVM.	Presents a comparative study and gets an accuracy of 95% with an SVM classifier.
[20] Prakash et all(2017)	GLCM,SVM,K-Means	The research might help diagnose various plant illnesses and increase the classification accuracy, which is now around 90%.
[21] Phadikar et al (2008)	Bayes and SVM classifier, mean filtering technique, and Otsu's algorithm	Accuracy : Baye's – 68.1 % SVM – 79.5% accuracy

From the survey done on the existing techniques, we found many techniques, where the most popular being K-means , SVM and Bayesian classification, and ANN(Artificial Neural Networks). We were not able to find many techniques using Transfer Learning for the purpose.

Based on the review, it was observed that though some work has been done in the field, no widespread work has been done taking into account a generalized and large number of plants and many diseases. Most of the present works are concentrated on diseases in 1 category of plants.

Therefore, it is required to propose a generalized approach to predict numerous diseases of several plants.

As a result, the goal of this research is to provide an approach/method for classifying leaf detection with improved accuracy rates for different leaves and disease categories.

3. Proposed Method

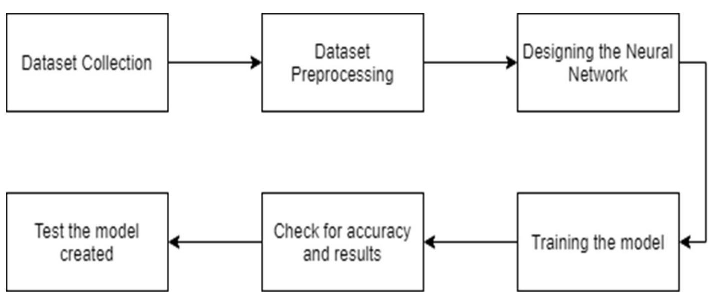

Figure 1: Flowchart showing the overall development of the method

3.1. Dataset Collection

For any supervised learning project, the main component is the dataset. For this project, we used the publicly available PlantVillage dataset [5] on Kaggle.

The dataset was accessed directly from Kaggle Notebook which was used to make the model for the paper.

Over 50,000 pictures of healthy and diseased plant leaves are included in this collection. It contains the infected images in 38 categories where 26 categories of diseases of the 12 plants, namely, tomato, apple blueberry, cherry, grapes, corn, orange, peach, raspberry, soya bean, squash, and strawberry Input leaf image.

The disease into which the diseases were classified:-

Table 2: Fruit Diseases

Serial Number	Fruit	Disease
1.	Apple	Apple scrab
		Apple Black rot
		Cedar apple rust
2.	Cherry	Cherry Powdery Mildew
3.	Corn	Corn Cercospora leaf spot Gray leaf spot
		Corn Common Rust
		Corn (maize) Northern Leaf Blight
4.	Grape	Grape Black Rot
		Grape Esca (Black Measles)
		Grape Esca (Black Measles)
		Grape Leaf blight (Isariopsis Leaf Spot)
5.	Orange	Orange Huanglongbing (Citrus greening)
6.	Peach	Peach Bacterial spot
7.	Pepper	Bell Bacterial spot
8.	Potato	Potato Early blight
		Potato Late blight
9.	Squash	Squash Powdery mildew
10.	Strawberry	Strawberry Leaf scorch
11.	Tomato	Tomato Bacterial spot
		Tomato Early blight
		Tomato Late blight
		Tomato Leaf Mold
		Tomato Septoria leaf spot
		Tomato Spider mites Two-spotted spider mite
		Tomato Target Spot

		Tomato Yellow Leaf Curl Virus
		Tomato mosaic virus

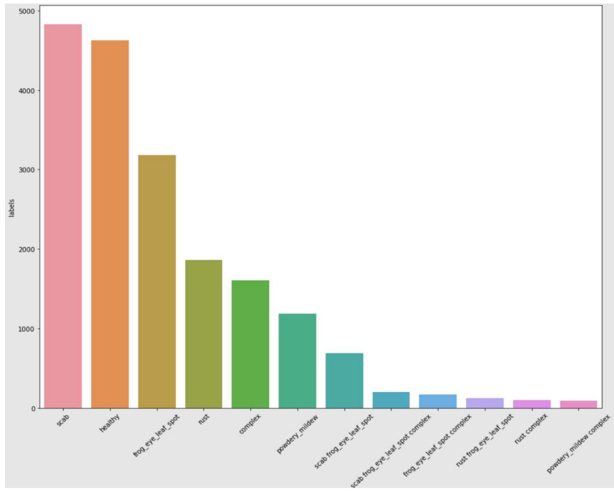

Figure 2: Distribution of pictures of leaves of different plants

3.2. Data Pre-Processing

For the dataset, the image contains background noise. In order to extract the relevant region from the input picture, the Tiramisu model must be used. It is based on DensNet,where all the layers are interconnected. Also the Tiramisu model adds skip connections to the up-sampling layer like Unet.

3.3. Designing the Neural Network

For making the model which would be trained on the images, we incorporated the transfer learning technique.

Transfer learning is a novel method to deep learning in which models that have been pre-trained for one task are repurposed for another.

One of the biggest advantages being that there is no need for extra feature extraction step. This is because they are very deep neural networks where the initial layers act as feature extractor.

In our case we went on with the use of Inception V3 pre-trained neural network. It is a family of Inception neural networks where all the previous features of inception v1 and v2 are incorporated along with label smoothing, factorized 7x7 optimizer, RMSProp optimizer and BatchNorm.

Also, in comparison with its counterparts like VGGNet, Inception networks work better and provide more computationally efficiency both in the terms of parameters generated by the network and the economical cost incurred in terms of memory and other resources.

In this model we added further layers at the end to get the prediction into the 5 categories as we desired. After the model being compiled with trained it with our large dataset with 25 epoch and batch size of 16.

3.4. Training the model

The model is being trained on the training dataset, which is produced by taking 70% of the entire dataset for training and 30% for testing.

3.5. Checking for Accuracy Achieved

In the model, we were able to achieve an accuracy of 96.74% on Validation Accuracy.

4. Experimental Results and Analysis

In this section, we are briefly explaining the result of the proposed model.

4.1. Results of Data Preprocessing

With the Tiramisu model, we were able to extract the only leaf image from the overall image.

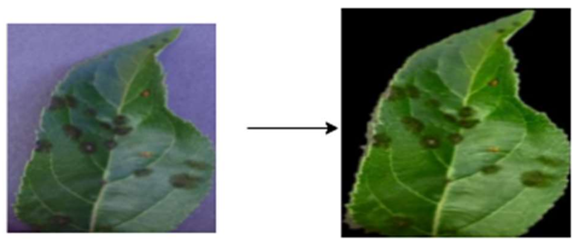

Figure 3: Image showing the segmented image of the leaf A: Original leaf image with the background. B: Image of leaf extracted and background removed.

4.2. Result of the proposed model

The approach aimed to capture a large base of leaves that can be checked for any diseases. With the work, we were able to achieve the aim of targeting 12 different plant leaves and detecting 26 different diseases.

With the work we were able to achieve pretty good performance as we can see from the visualizations.

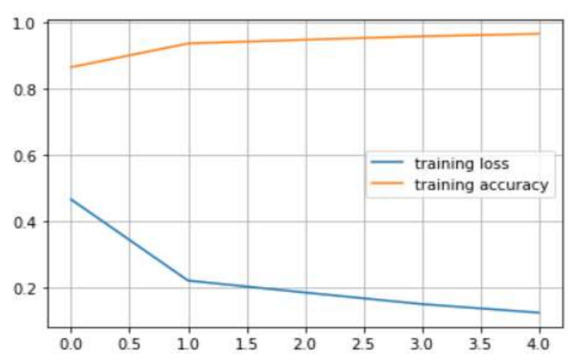

Figure 4: Graph showing the accuracy of the training phase

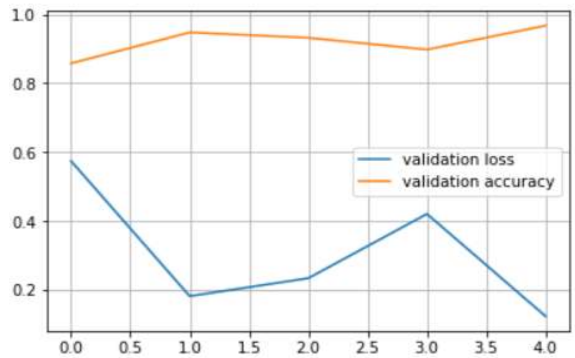

Figure 5: Graph showing the accuracy of the validation phase

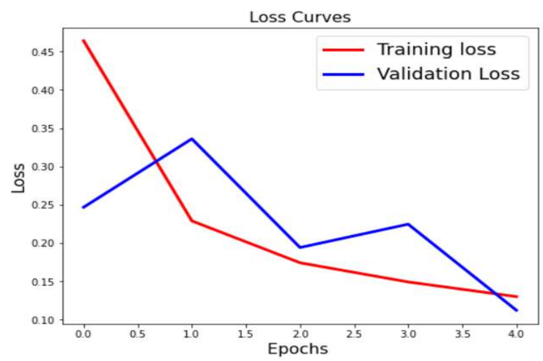

Figure 6: Graph showing the loss of the validation phase and training phase

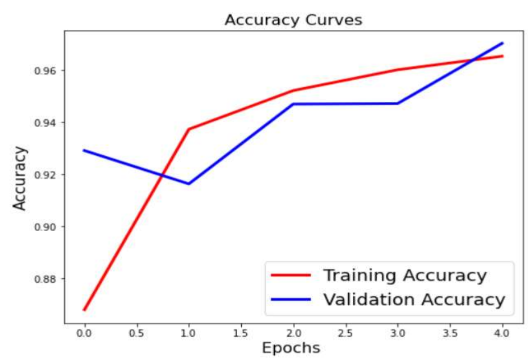

Figure 7: A graph depicting the validation and training phases' accuracy.

The system when fed with the input of the image of a plant leaf, predicts the top 5 diseases with bar chart visualization. With this, we aim to address anomaly that can occur in predictions. Meaning it provides options and its opinion on the diseases that the plant leaf has with the amount of confidence in each diseases.

Also it provides the user,i.e., farmers to view alternative diseases that can be present.

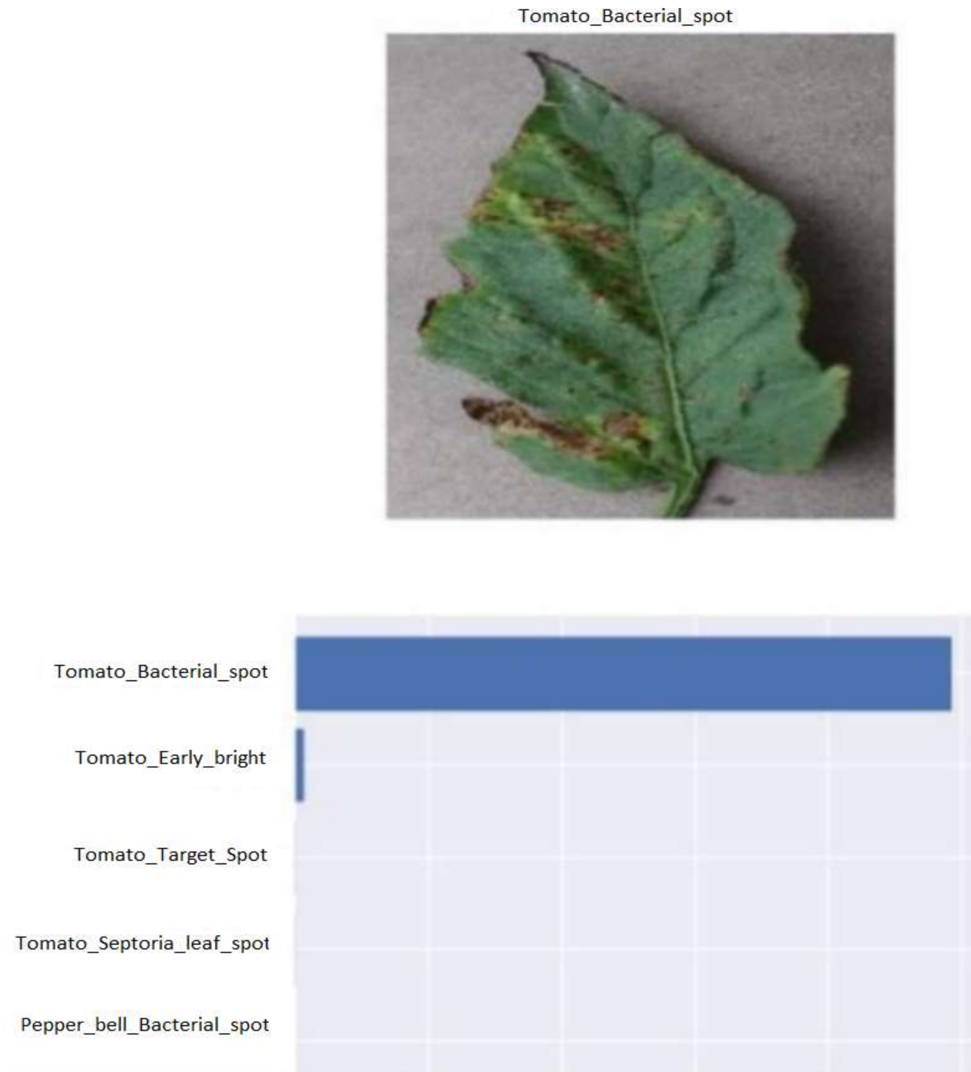

Figure 8: Image showing the prediction of the model on an input image

5. Conclusion

We had started work on the paper focusing on detecting multiple diseases in multiple plants. We focused on 12 categories of plants namely-

Table 3: Fruit category

Tomato	Apple	Blueberry	Cherry
Grapes	Corn	Orange	Peach
Raspberry	Soya Bean	Squash	Strawberry

On our model, we achieved an accuracy of 96.74%. With our paper, we conclude that with the model developed we can classify the images into vast 26 categories of the diseases from 11 different plant types efficiently and effectively with an overall accuracy of 96.74% higher than any of the existing works and also into more categories than any existing works.

We think that the future work on this can be to convert this model into an application, which can easily be used by the real farmers to use the model for fast and efficient disease identification.

6. References

[1] Ghaiwat, S. N., & Arora, P. (2014). Detection and classification of plant leaf diseases using image processing techniques: a review. International Journal of Recent Advances in Engineering & Technology, 2(3), 1-7.

[2] Dhaygude, S. B., & Kumbhar, N. P. (2013). Agricultural plant leaf disease detection using image processing. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 2(1), 599-602.

[3] Badnakhe, M. R., & Deshmukh, P. R. (2011). An application of K-means clustering and artificial intelligence in pattern recognition for crop diseases. In International conference on advancements in information technology (Vol. 20, pp. 134-138).

[4] Arivazhagan, S., Shebiah, R. N., Ananthi, S., & Varthini, S. V. (2013). Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. Agricultural Engineering International: CIGR Journal, 15(1), 211-217.

[5] Kulkarni, A. H., & Patil, A. (2012). Applying image processing technique to detect plant diseases. International Journal of Modern Engineering Research, 2(5), 3661-3664.

[6] Bashir, S., & Sharma, N. (2012). Remote area plant disease detection using image processing. IOSR Journal of Electronics and Communication Engineering, 2(6), 31-34.

[7] Naikwadi, S., & Amoda, N. (2013). Advances in image processing for detection of plant diseases. International Journal of Application or Innovation in Engineering & Management, 2(11).

[8] Patil, S. B., & Bodhe, S. K. (2011). Leaf disease severity measurement using image processing. International Journal of Engineering and Technology, 3(5), 297-301.

[9] Chaudhary, P., Chaudhari, A. K., Cheeran, A. N., & Godara, S. (2012). Color transform based approach for disease spot detection on plant leaf. International journal of computer science and telecommunications, 3(6), 65-70.

[10] Rathod, A. N., Tanawal, B., & Shah, V. (2013). Image processing techniques for detection of leaf disease. International Journal of Advanced Research in Computer Science and Software Engineering, 3(11).

[11] Mondal, D., Chakraborty, A., Kole, D. K., & Majumder, D. D. (2015, October). Detection and classification technique of Yellow Vein Mosaic Virus disease in okra leaf images using leaf vein extraction and Naive Bayesian classifier. In 2015 International Conference on Soft Computing Techniques and Implementations (ICSCTI) (pp. 166-171). IEEE.

[12] Padol, P. B., & Yadav, A. A. (2016, June). SVM classifier based grape leaf disease detection. In 2016 Conference on advances in signal processing (CASP) (pp. 175-179). IEEE.

[13] Reza, Z. N., Nuzhat, F., Mahsa, N. A., & Ali, M. H. (2016, September). Detecting jute plant disease using image processing and machine learning. In 2016 3rd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT) (pp. 1-6). IEEE.

[14] Tejonidhi, M. R., Nanjesh, B. R., Math, J. G., & D'sa, A. G. (2016, March). Plant disease analysis using histogram matching based on Bhattacharya's distance calculation. In 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) (pp. 1546-1549). IEEE.

[15] Arya, M. S., Anjali, K., & Unni, D. (2018, January). Detection of unhealthy plant leaves using image processing and genetic algorithm with Arduino. In 2018 International Conference on Power, Signals, Control and Computation (EPSCICON) (pp. 1-5). IEEE.

[16] Mehra, T., Kumar, V., & Gupta, P. (2016, December). Maturity and disease detection in tomato using computer vision. In 2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC) (pp. 399-403). IEEE.

[17] Pawar, P., Turkar, V., & Patil, P. (2016, August). Cucumber disease detection using artificial neural network. In 2016 International Conference on Inventive Computation Technologies (ICICT) (Vol. 3, pp. 1-5). IEEE.
[18] Narmadha, R. P., & Arulvadivu, G. (2017, January). Detection and measurement of paddy leaf disease symptoms using image processing. In 2017 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-4). IEEE.
[19] Tripathi, M. K., & Maktedar, D. D. (2016, August). Recent machine learning based approaches for disease detection and classification of agricultural products. In 2016 International Conference on Computing Communication Control and automation (ICCUBEA) (pp. 1-6). IEEE.
[20] Prakash, R. M., Saraswathy, G. P., Ramalakshmi, G., Mangaleswari, K. H., & Kaviya, T. (2017, March). Detection of leaf diseases and classification using digital image processing. In 2017 international conference on innovations in information, embedded and communication systems (ICIIECS) (pp. 1-4). IEEE.
[21] Phadikar, S., & Sil, J. (2008, December). Rice disease identification using pattern recognition techniques. In 2008 11th International Conference on Computer and Information Technology (pp. 420-423). IEEE.
[22] Gurjar, A. A., & Gulhane, V. A. (2012). Disease detection on cotton leaves by eigen feature regularization and extraction technique. International Journal of Electronics, Communication and Soft Computing Science & Engineering (IJECSCSE), 1(1), 1.
[23] Al Bashish, D., Braik, M., & Bani-Ahmad, S. (2010, December). A framework for detection and classification of plant leaf and stem diseases. In 2010 international conference on signal and image processing (pp. 113-118). IEEE.
[24] Wu, S. G., Bao, F. S., Xu, E. Y., Wang, Y. X., Chang, Y. F., & Xiang, Q. L. (2007, December). A leaf recognition algorithm for plant classification using a probabilistic neural network. In 2007 IEEE international symposium on signal processing and information technology (pp. 11-16). IEEE.
[25] https://www.kaggle.com/abdallahalidev/plantvillage-dataset