Deep Convolutional Neural Network for Real Time Visual Inspection of Structural Audit

Aditya Shenoy – 161080036
Rahul Bhoir – 161080063
Kalyani Borkar – 171081990
Rani Deore – 171081993

Abstract

Key Points

- Structural audit requires manual expertise of surveyors that can be subjective and erroneous
- Digital Image Processing needs to be hardcoded and Deep Learning has been applied for only crack detection
- There is a need for a single generalized model, which can classify more than one structural defect during the visual inspection of structural audit

Key Points (Cont.)

- The proposed system classifies cracks, dampness and paint peel-off using a single generalized model as a part of visual inspection while doing structural audit of infrastructures
- The contribution of this work is 95% overall accuracy, 95% average precision, 95.25% average recall, and 95% average F1-score
- Usage of a non-complex model has reduced the time taken for training the model and getting predictions from the model

Introduction

Background

- Construction of different infrastructures is a necessity for human life
- Old buildings show cracks, paint peel off, and dampness
- Buildings need to be audited after 30 years
- Visual inspection for detection of cracks, dampness, paint peel off totally depends on the surveyor’s experience and knowledge
- There may be human errors

Literature Review

Digital Image Processing

- MESG algorithm is use for enhancing the noisy image
- SOBEL filter is used for obtaining binary image
- CANNY edge detection is used for feature extraction
- HMRF-EM algorithm is used to increase the accuracy to label the data
- Particularly good accuracy for detection crack (up to 90% accuracy)
- Numerous algorithm available for usage: grouping different algorithms can be used to detect crack.
- Grouping of Algorithms for one problem cannot be used for another problem

Deep Learning for Crack Detection

- Deep Convolutional Neural Network (DCNN) is capable to sense the crack robustly while dealing with complicated background images
- Hessain matrix based linear filtering approach used to increase crack area and regulate threshold gap to get crack binarization segmentation result.
- Using Internet of Things (IoT) technology for collecting video data is easy and solves the problems of high risk factor in domestic and low fracture analysis.
- Genetic Algorithm for CNN model evolved several parameters that dictate the structure of CNN

Structural Audit

- Overall health examination of a building to ensure its safety
- It identifies parts of a building that may be in need of immediate repair, renovation or replacement
- Should be conducted once in 5 years for 15-30 years old buildings
- Should be conducted once every 3 years for 30+ years old buildings
- A structural audit is a highly recommended preventive measure to avoid any calamitous eventualities altogether

Deep Learning on Video Images

- Using Full Convolution Network (FCN) enables pixel level classification
- Temporal features of subsequent frames can be captured using Convolutional Long Short Term Memory (Conv-LSTM)
- Faster Region-based Convolutional Neural Network (R-FCN) gives better accuracy results
- You Look Only Once (YOLO) processes the input faster and can be used in real time classification

Problem Statement

There is a need for a single generalized model, which can classify more than one structural defect during the visual inspection of structural audit

Proposed Solution

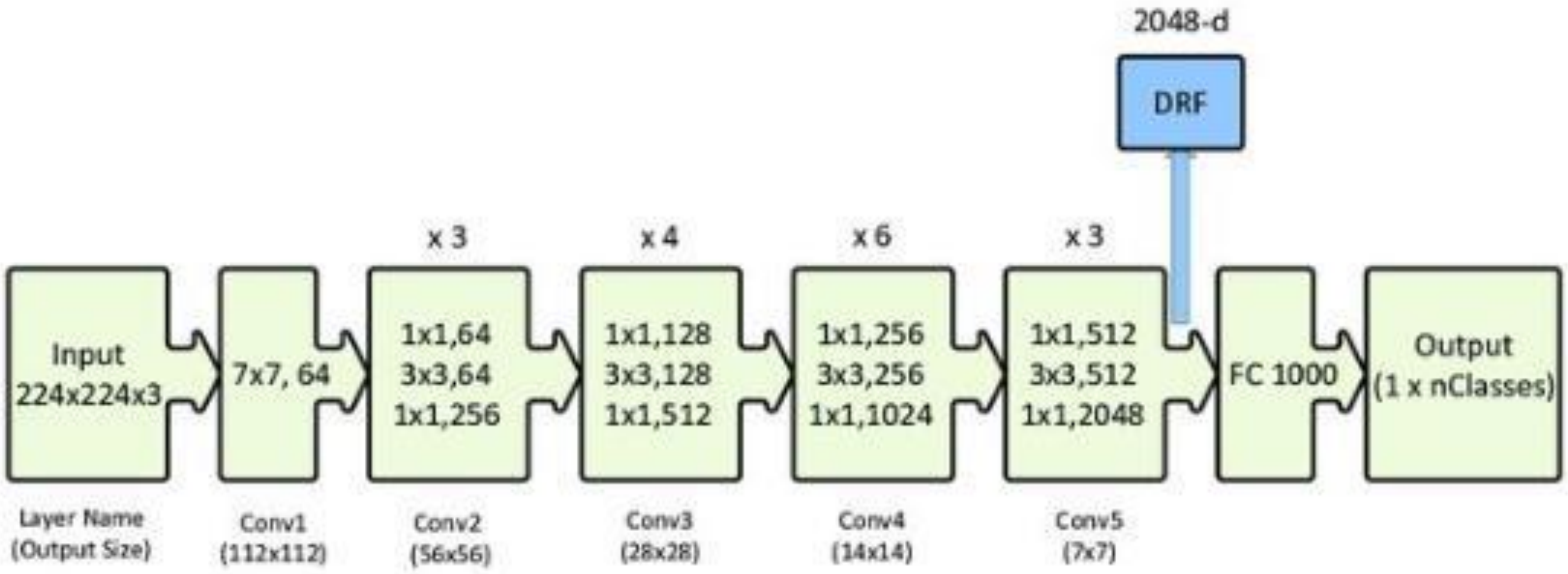
Proposed System

The proposed system classifies cracks, dampness and paint peel-off using a single generalized model as a part of visual inspection while doing structural audit of infrastructures

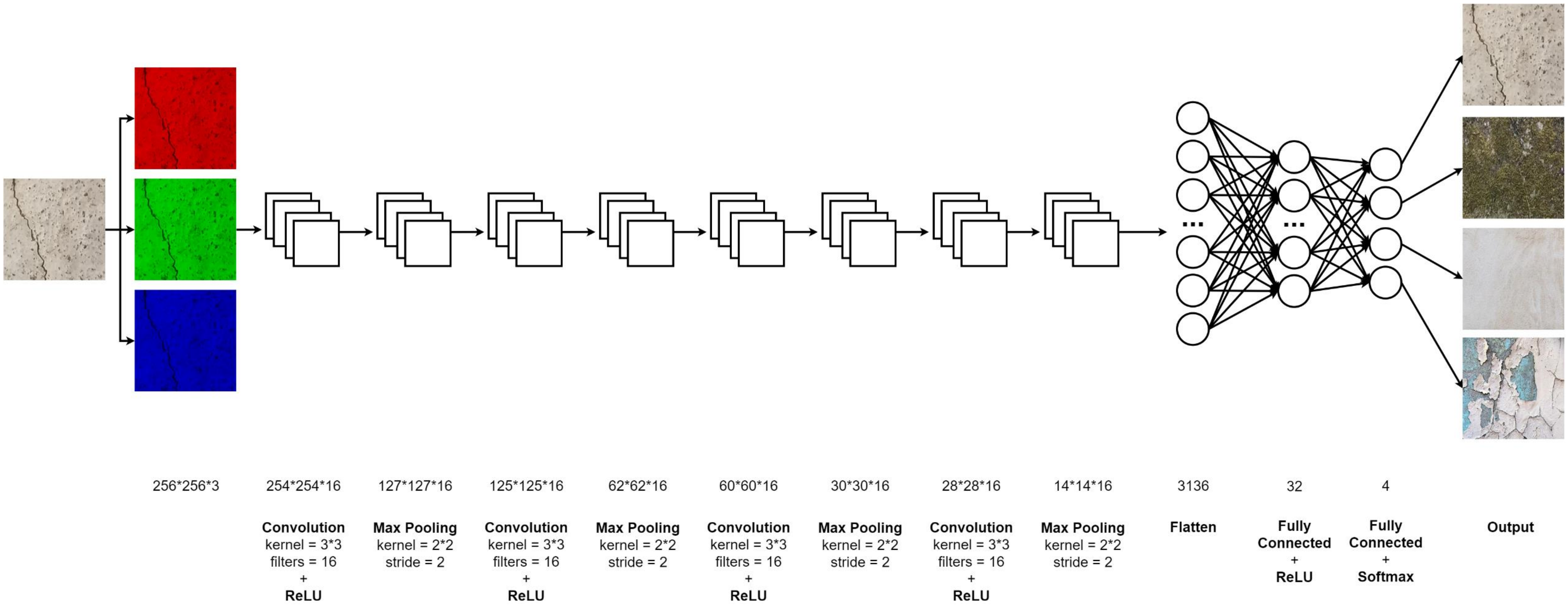
Transfer Learning

Transfer learning is a research problem in machine learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem

ResNet-50 Model Architecture



Effective Defect Classifier (EDC)



Sliding Window Mechanism

- The video is divided into frames and each frame is divided into grids
- Each sub-image in the grid is passed to the model for prediction
- The predictions are then color coded
- The color coded result is pasted along with the original video frames

Results and Analysis

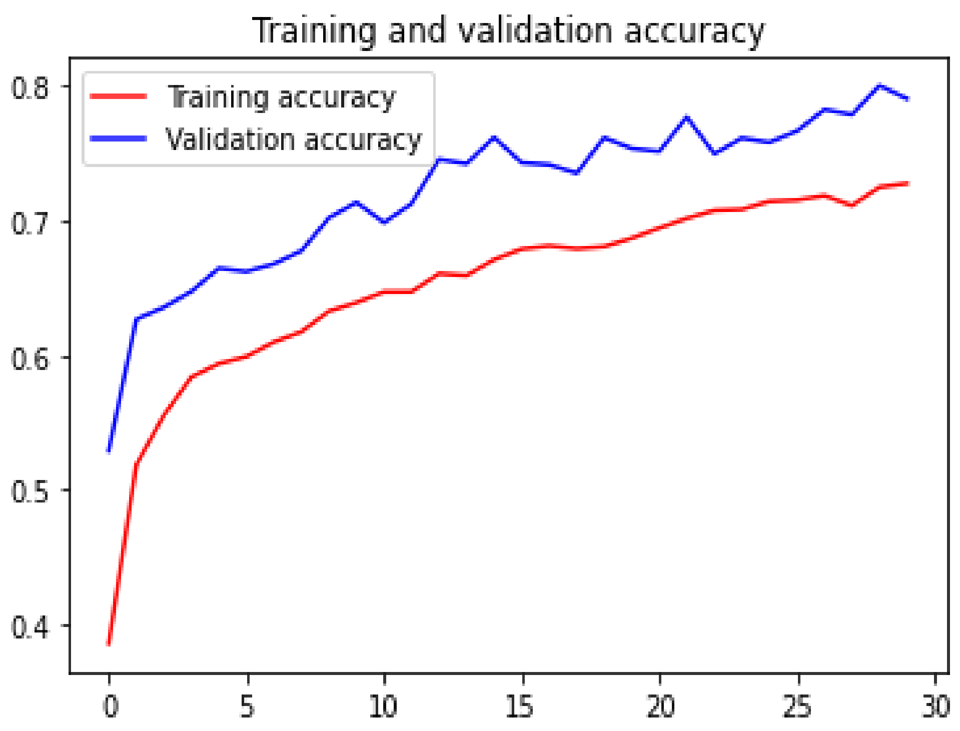
Dataset

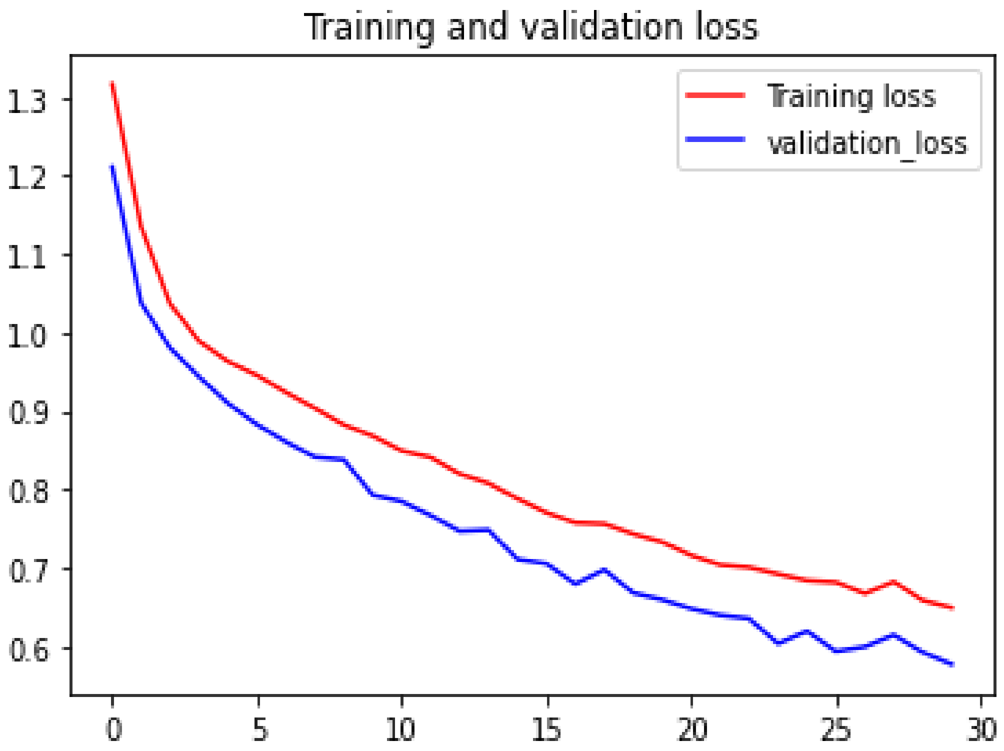
- 4 Classes – Crack, Dampness, Paint Peel Off, Plain Wall
- Image Dimensions – 256*256
- Training Size – 32,000
- Testing Size – 4,000
- Non Distorting Transformations – 7
- Distorting Transformations – 2

Transfer Learning

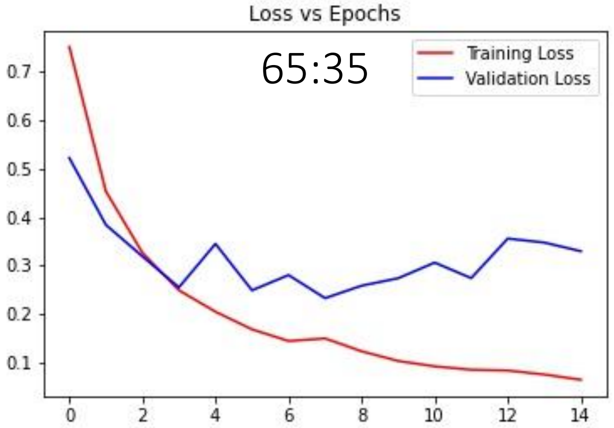
- ResNet-50 is not similar to our problem
- Validation accuracy of the model trained was very poor
- 9 out of 10 images were incorrectly classified.

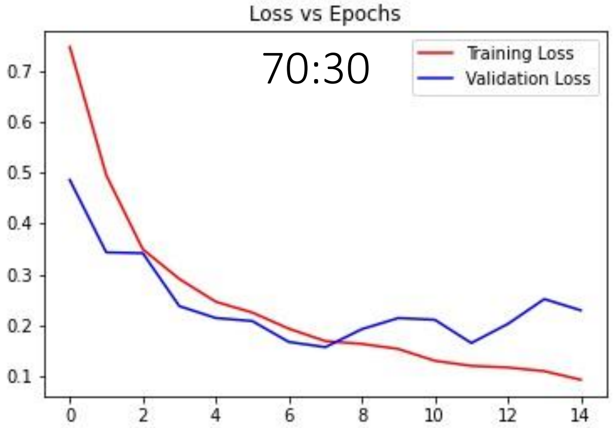
Transfer Learning Model Accuracy and Loss

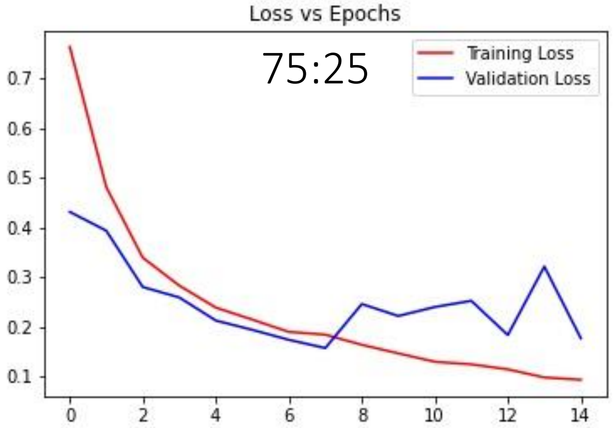


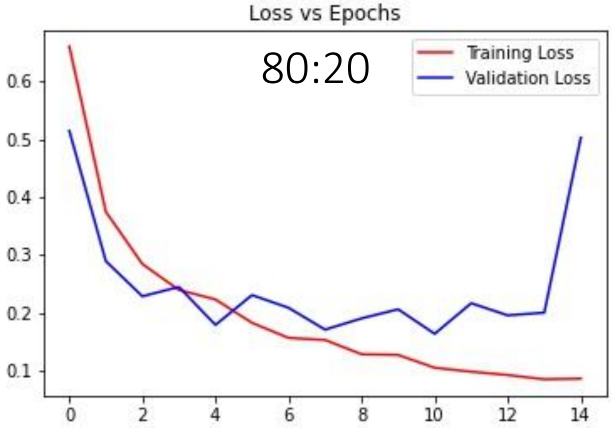


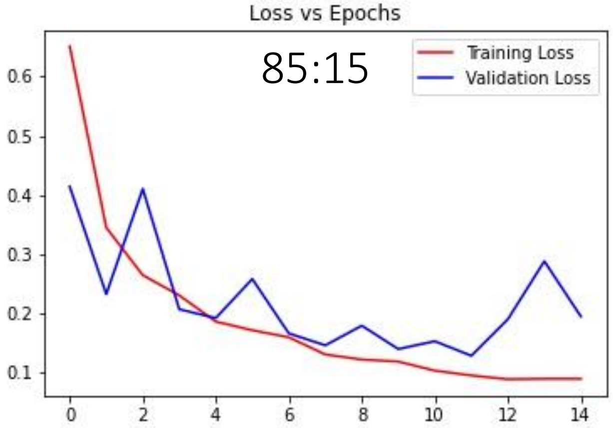
EDC - Train and Test Split Ratio

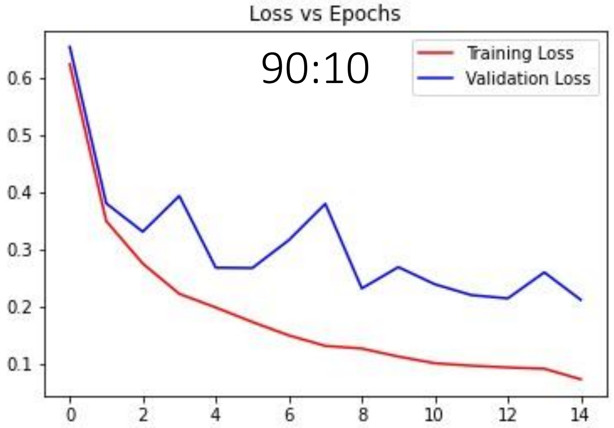












EDC – Confusion Matrix

Predicted→ Actual↓	C	D	N	P	Total
C	932	27	13	28	1000
D	16	966	8	10	1000
N	8	10	957	25	1000
P	24	11	19	946	1000

	C	Not C
C	TP = 932	FN = 68
Not C	FP = 48	TN = 2952

	D	Not D
D	TP = 966	FN = 34
Not D	FP = 48	TN = 2952

	N	Not N
N	TP = 957	FN = 43
Not N	FP = 40	TN = 2960

	P	Not P
P	TP = 946	FN = 54
Not P	FP = 43	TN = 2957

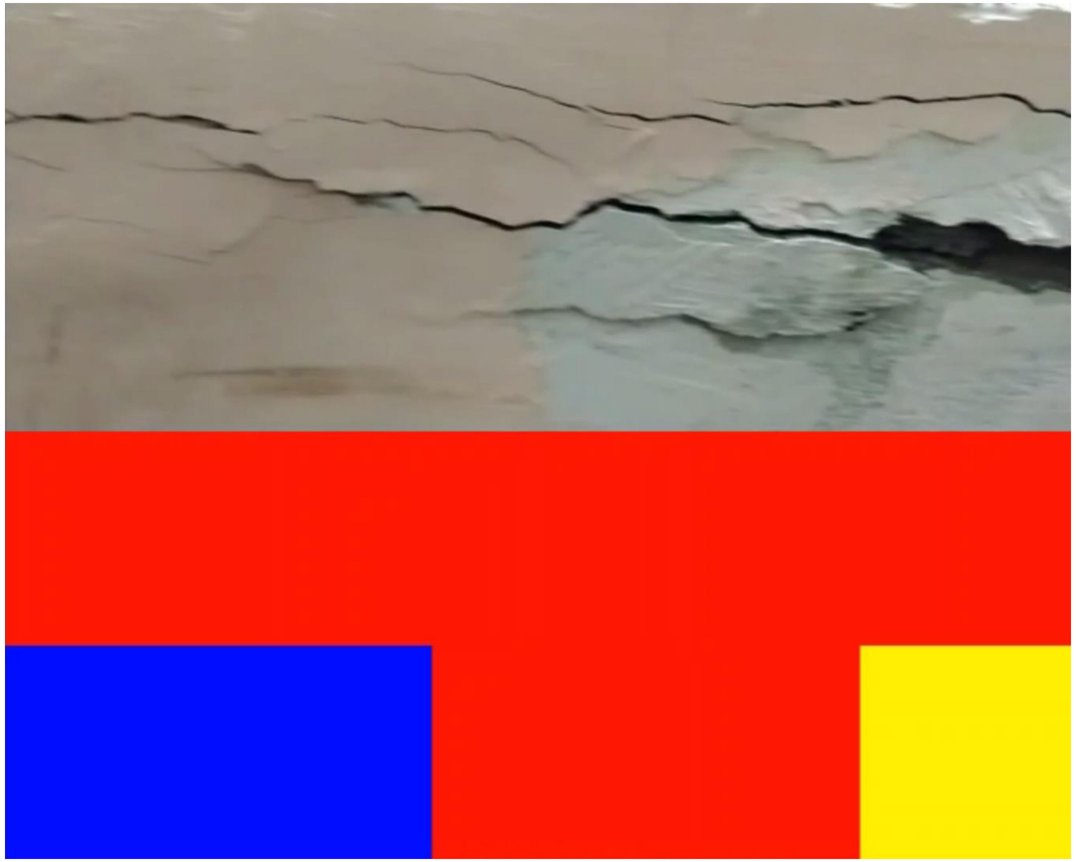
C – Crack | D – Dampness | N – Plain Wall | P – Paint Peel Off

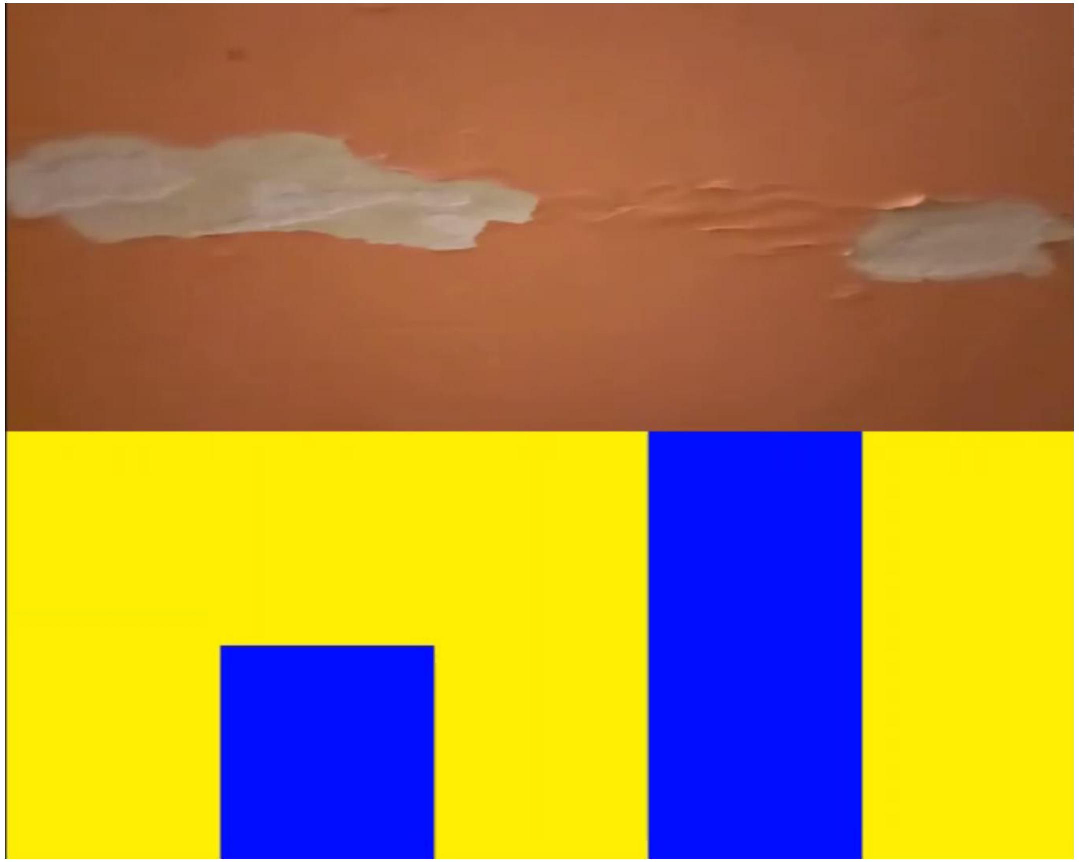
EDC – Metrics

- Training Time – 77.6 seconds per epoch on an average
- Accuracy – 95%

	Precision	Recall	F1-Score	Total
Crack	0.95	0.93	0.94	1000
Dampness	0.95	0.97	0.96	1000
Plain Wall	0.96	0.96	0.96	1000
Paint Peel Off	0.94	0.95	0.94	1000

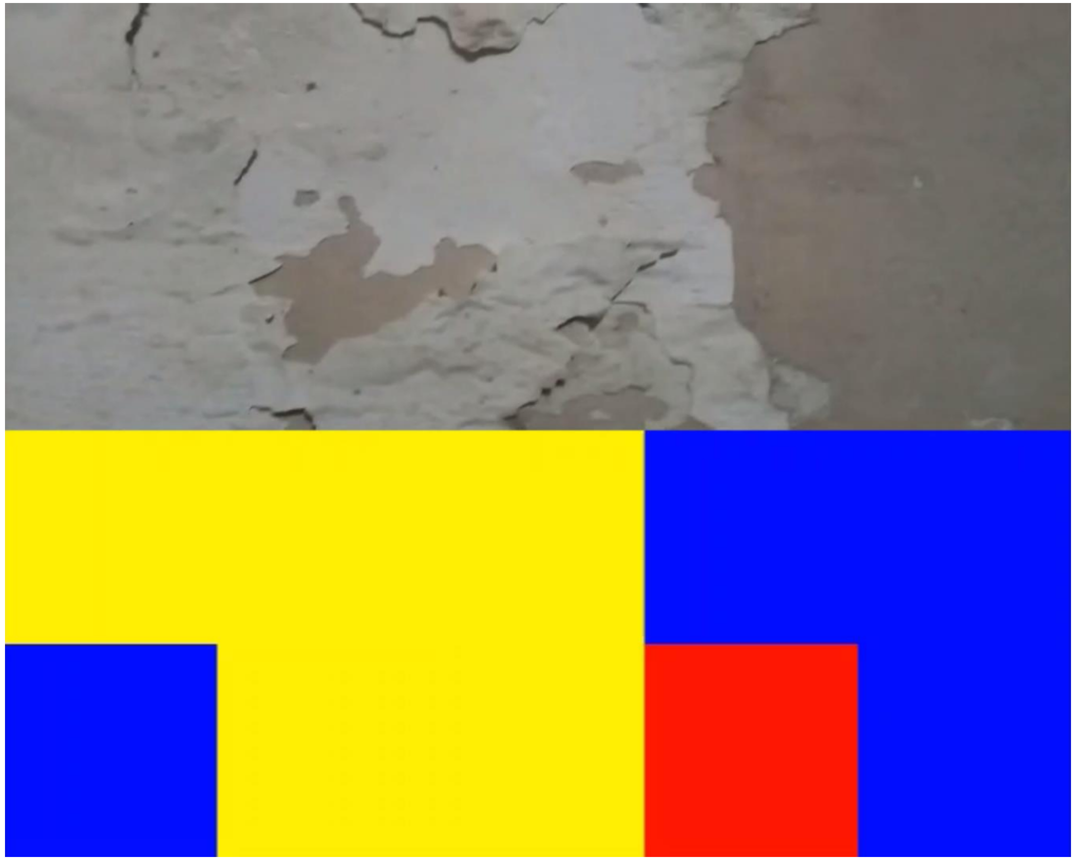
Sliding Window Output Frames





Red – Crack | Green – Dampness | Blue – Plain Wall | Yellow – Paint Peel Off

Sliding Window Output Frames Cont.



Red – Crack | Green – Dampness | Blue – Plain Wall | Yellow – Paint Peel Off

Conclusion and Future Scope

Conclusion

- EDC classifies more labels than existing models
- Use of non complex model has reduced training and prediction times
- Average metric for Accuracy, Precision, Recall, and F1-Score is 95%

Future Scope

- Dataset size can be increased by doing more surveys
- More types of defects can be incorporated into the model
- Pre-processing techniques can be applied to get more features from images

Appendix

Training Time History

Epochs	Training Time (sec)	Epochs	Training Time (sec)
1	80	9	77
2	80	10	77
3	78	11	77
4	78	12	77
5	78	13	77
6	77	14	77
7	77	15	77
8	77	-	-
Avg. Training Time per Epoch		77.6 seconds	
Total Training Time for 15 Epochs		19 minutes 24 seconds	

Accuracy and Loss History

Epochs	Training Accuracy (%)	Validation Accuracy (%)	Training Loss	Validation Loss
1	70.36	71.95	0.71	0.70
2	85.69	91.30	0.39	0.35
3	90.88	92.90	0.29	0.22
4	92.09	90.68	0.25	0.30
5	92.60	92.37	0.21	0.22
6	93.79	91.15	0.19	0.28
7	93.74	93.15	0.17	0.21
8	95.06	93.48	0.17	0.23

Accuracy and Loss History (Cont.)

Epochs	Training Accuracy (%)	Validation Accuracy (%)	Training Loss	Validation Loss
9	95.06	94.67	0.13	0.17
10	95.12	94.35	0.13	0.24
11	95.51	93.48	0.12	0.27
12	96.16	94.23	0.10	0.20
13	96.57	95.15	0.09	0.23
14	96.41	94.92	0.10	0.24
15	97.12	95.07	0.08	0.26