**Visualizing Uncertainty with Chromatic Aberration**

By

Md Rashidul Islam  
md313724@dal.ca

Supervised by

Dr. Stephen Brooks

Professor

Faculty of Computer Science, Dalhousie University

Submitted in partial fulfilment of the requirements

for the degree of Master of Computer Science

at

Dalhousie University

Faculty of Computer Science, Dalhousie University

Halifax, Nova Scotia

© Dalhousie University 2020. All rights reserved.

**Table of Contents**

List of Tables ………………………………………………………………………………….. vii

List of Figures …………………………………………………………………………………. viii

List of Algorithms …………………………………………………………………………….. x

Abstract ………………………………………………………………………………………... xi

List of abbreviations ……………………………………………………………………...…... xii

Acknowledgements ………………………………………………………………………….. xiii

**Chapter 1 Introduction** …………………………………………………………………... 1

1.1 Background and Motivation ……………………………………………………….... 1

1.2 Background Concepts ……………………..………………………………………… 2

1.2.1 Machine Learning (Predictive Models) …………………………………... 2

1.2.2 D3.js ……………………………………………………………................. 3

1.2.3 Uncertainty …………………………………………………...…………… 3

1.2.4 Streamgraph ……………………………………………………………….. 4

1.2.5 Texture …………………………………………………………………….. 5

1.2.6 Chromatic Aberration …………………………….……………………….. 6

1.3 Problem Statement …………………………………………………………………. 7

1.4 Approach ……………………………………………………………………………. 8

1.5 Thesis Outline ……………………………………………………………………….. 9

**Chapter 2 Literature Review** …………………………………………………………… 10

2.1 Introduction ………………………………………………………………………… 10

2.2 Prior works related to prediction in Machine Learning Models …………………… 10

2.3 Uncertainty related prior works ……………………………………………………. 11

2.4 Chromatic Aberration related prior works …………………………………………. 19

2.5 Texture related prior works ………………………………………………………… 21

2.6 Evaluation of Visualization Systems ……………………………………………….. 23

2.7 Limitations of related works ……………………………………………………….. 26

**Chapter 3 Data Collection, Processing, and Introduction of Models** ……………… 29

3.1 Introduction ……………………………………………………………………….. 29

3.2 Data ……………………………………………………………………………….. 29

3.2.1 Data Collection …………………………………………………………. 29

3.2.2 Sample Data ……………………………………………………………. 31

3.3 Machine Learning Algorithms …………………………………………………… 31

3.3.1 Predictive/Forecasting Models …………………………………………. 32

3.3.2 Time Series Analysis vs Forecasting …………………………………… 32

3.3.3 Concerns of Forecasting ………………………………………………… 33

3.3.4 Example of Forecasting …………………………………………………. 34

3.4 MLP ……………………………………………………………………………….. 34

3.5 CNN ……………………………………………………………………………….. 38

3.6 LSTM ……………………………………………………………………………… 40

3.7 ARIMA ……………………………………………………………………………. 41

3.7.1 Auto ARIMA ……………………………………………………………. 42

3.8 Uncertainty Data Generation ……………………………………………………… 43

3.8.1 Uncertainty Data Scaling ………………………………………………... 44

3.8.2 Snapshot of uncertainty data ……………………………………………. 45

3.8.3 Top 10 uncertainty countries using MLP model ………………………… 45

3.8.4 Lowest 10 uncertainty countries using MLP model …………………….. 45

3.8.5 Uncertainty Comparison among Models ………………………………… 46

**Chapter 4 Visualization Method** ……………..………………………………………. 47

4.1 Introduction ………………………………………………………………………. 47

4.2 Background Architecture ………………………………………………………… 47

4.3 Examples of CA in Shapes ………………………………………………………. 48

4.4 Texture Pattern Generation ………………………………………………………. 49

4.4.1 Slicing Plot …………………………………………………………….. 51

4.4.2 Pattern Generation ……………………………………………………… 52

4.4.3 Texture Generation …………………………………………………….. 54

**Chapter 5 Application of Chromatic Aberration** ……………..…………………… 55

5.1 Introduction ………………………………………………………………………. 55

5.2 Web Interface ……………………………………………………………………. 55

5.3 Filtering ………………………………………………………………………….. 57

5.3.1 Bubble Selection Mode ……………………………………………….. 57

5.3.2 Bubble Removal Mode ……………………………………………….. 58

5.4 Legend ……………………………………………………………………………. 58

5.5 Reshuffling Streamgraph …………………………………………………………. 58

5.6 Drill-down All Model Predictions ……………………………………………….. 60

5.7 Star Fish Inspired Design …………………………………………………………. 61

5.8 Parallel Coordinate Chart …………………………………………………………. 63

5.9 Bubble Grid Chart ………………………………………………………………… 64

5.10 Horizontal Chart …………………………………………………………………. 65

5.11 Square Grid Chart ……………………………………………………………….. 67

5.12 World Map ………………………………………………………………………. 68

**Chapter 6 Evaluation: User Study Design** …….…………………………………… 70

6.1 Introduction …………………………………………………………………………70

6.2 Background and Goal ……………………………………………………………… 70

6.3 Research Questions ……………………………………………………………….... 71

6.4 Study Material …………………………………………………………………….... 71

6.4.1 Technology and Browser Support………...…………………………….... 71

6.4.2 Study Components………...……………………………........................... 72

6.4.3 Counter Balancing………...……………………………........................... 73

6.5 Recruitment ……………………………………………………………………….... 75

6.5.1 Criteria ………………………………………………………………….... 75

6.5.1 Hiring Procedure ……………………………………………………….... 76

6.5.3 Making Schedule ……………………………………………………….... 76

6.6 Study Procedure .………………………………………………………………….... 76

6.6.1 Start Event …….……………………………………………………….... 77

6.6.2 Briefing ….…….……………………………………………………….... 77

6.6.3 Color Blindness Test ….…….………………………………………….... 77

6.6.4 Pre-Session Discussion ……..………………………………………….... 78

6.6.5 Questionnaire ……...….…….………………………………………….... 78

6.7 Data Collection and Storing …….…….………………………………………….... 77

6.8 Session Ending …………...…….…….…………………………….…………….... 77

**Chapter 7 Evaluation: Results and Numerical Analysis**…….……………………... 83

7.1 Introduction ……………………………………………………………………….. 83

7.1.1 Sample Population Demographics ………………..…………………….. 83

7.2 Study Results ………………………..…………………………………………….. 84

7.2.1 Quantitative Questionnaire Results ………………….………………….. 84

7.2.1.1 One-way repeated measures ANOVA ……………….……….. 84

7.2.1.2 Paired t-test …………………………………………………… 88

7.2.2 Time Utilization Results …………………………….………………….. 91

7.2.3 SUS Results ………………………………………….………………….. 93

7.2.4 NASA-TLX Results ………………………………….………………….. 95

7.3 User Comments …………………..………………………………………………... 97

7.2 Summary of the Results ……………..…………………………………………….. 97

**Chapter 8 Conclusion and Future Work** ………………………….………………… 98

**References** ………………………….……………………………………………………… 99

Appendix A REB Application Form ……………………………………………………… 105

Appendix B Informed Consent Form …………….………………………………………. 121

Appendix C First Email …………….…………………..…………………………………. 124

Appendix D Second Email …………….………………………….………………………. 125

Appendix E Examples and Questionnaire ………………………………………………. 126

E.1 Questionnaire Setup and Arrangement ………………………………………. 126

E.2 Example of CA + Bubble ……………………………………………………. 130

E.3 Questionnaire on CA + Bubble ……………………………………………… 131

E.4 Example of VSUP + Bubble ………………………………………………… 133

E.5 Questionnaire on VSUP + Bubble …………………………………………... 134

E.6 Example of CA + Grid ……………………………………………………… 136

E.7 Questionnaire on CA + Grid ………………………………………………… 137

E.8 Example of VSUP + Grid …………………………………………………… 139

E.9 Questionnaire on VSUP + Grid ……………………………………………... 140

E.10 Questions on System Usability Scale (SUS) …………………………….. 142

E.11 Questions on NASA TLX ……………………………………………….. 143

Appendix F Ishihara Color Blindness Test Plates ……………………………………… 144

Appendix G Amazon Gift Card Sample ………………………………………………… 150

Appendix H REB Approval Letter ……………………………………………………… 151

Appendix I User Study Results ………………………………………………………… 152

**List of Tables**

Table 3.1 COVID Data property list ………………………………………… 30

Table 3.2 Screenshot of sample data ………………………………………… 31

Table 3.3Top uncertainty countries ………………….……………………… 45

Table 3.4 Lowest Uncertainty countries …………………………………….. 45

Table 3.5 Uncertainty Comparisons of Models ……………………………… 46

Table 6.1 Task Arrangement of user study …………………………………... 74

Table 7.1 ANOVA Data Summary ………………………………………..… 84

Table 7.2 Shapiro-Wilk Normality Test for ANOVA ……………………….. 85

Table 7.3 ANOVA Test Results Summary ………………………………….. 86

Table 7.4 Summary of CA vs VSUP performance…….…………………….. 88 Table 7.5 Shapiro-Wilk Normality Test for t-test ………..………………….. 88 Table 7.6 Summary of CA vs VSUP timing ……………..………………….. 91 Table 7.7 SUS scores summary of CA vs VSUP ………..………………….. 93 Table 7.8 Normality test results of NASA-TLX score ….………………….. 95 Table 7.9 Kruskal-Wallis test results of NASA-TLX …...………………….. 96

Table I.1: Questionnaire Raw Scores of four components …..………………………152

Table I.2: Questionnaire Raw Scores of CA vs VSUP …………………………....…153

Table I.3: SUS Raw Scores for CA ………………………….………...…………..…154

Table I.4: SUS Raw Scores for VSUP ……………………….………..…………..…154

Table I.5: NASA-TLX Raw Scores for CA ……………….…………..…………..…155

Table I.6: NASA-TLX Raw Scores for VSUP …………….……….....…………..…156

Table I.7 Time Utilization for Full Questionnaire ……………..…………………… 158

**List of Figures**

Figure 1.1 Streamgraph Prototype ………………………….………………… 5

Figure 1.2 Example of Chromatic Aberration ……………………………….. 6

Figure 3.1 Predictive modelling workflow to generate uncertainty ………… 32

Figure 3.2 Example of daily covid forecasting for 200 days ……………….. 34

Figure 3.3 Basic Architecture of MLP network ……………………………... 35

Figure 3.4 Basic Architecture of CNN network ……………………………... 38

Figure 3.5 Basic Architecture of LSTM network ……………………………. 40

Figure 4.1 Underlying Geometry of Chromatic Aberration ………………… 48

Figure 4.2 Example CA on Bubbles and Rectangles ………………………… 49

Figure 4.3 Streamgraph Color Filled (left), Texture Filled (right) …………… 51

Figure 4.4 Sliced Streamgraph ……………………………………………….. 52

Figure 5.1 Initial Web Interface ……………………………………………… 55

Figure 5.2 Filter by selected countries of interest …………………………… 57

Figure 5.3 Removal of countries of interest ………………………………….. 58

Figure 5.4 Reshuffling Main Streamgraph …………………………………… 59

Figure 5.5 Streamgraphs for all models for Brazil …………………………… 60

Figure 5.6 Uncertainty presentation on stream by texture ……………………. 61

Figure 5.7 Multi Country Stream Graphs filled by color ……………………... 62

Figure 5.8 Multi Country Stream Graphs filled by Texture of CA …………… 63

Figure 5.9 Parallel coordinates chart ………………………………………….. 64

Figure 5.10 Bubble Grid chart with CA textures ………………………………. 65

Figure 5.11 Horizontal chart (Color filled) …………………………………….. 66

Figure 5.12 Horizontal chart (Texture filled) …………………………………... 66

Figure 5.13 Charts of Daily Counts for different countries ……………………. 67

Figure 5.14 Uncertainty in World view ………………………………………... 68

Figure 5.15 Zoomed World Map centering Nigeria …………………………… 69

Figure 6.1 Latin Square (left), Balanced Latin Square (right) ……………….. 73

Figure 6.2 Example Color Plate in Portal ……………………………………. 78

Figure 6.3 Flow of Questionnaires for a participant …………………………. 79

Figure 6.4 System Usability Scale Question Example………...……………… 79

Figure 6.5 NASA-TLX Work-Load Scale Question Example…...…………… 80

Figure 6.6 Session Ending Greetings…...…………………………………..… 82

Figure 7.1 Box plot of user performance …...…….………………………..… 85

Figure 7.2 Normal Distributions for components ...………………………..… 86

Figure 7.3 ANOVA Results...…………………………………..…………..… 87

Figure 7.4 Normal Distributions CA vs VSUP...…...…………..…………..… 87

Figure 7.5 Paired t-test gaussian plot for CA vs VSUP...…...………….......… 90

Figure 7.6 Paired t-test gaussian plot for time utilization...…...…..………..… 91

Figure 7.7 SUS rating plots for visualization methods ………………………. 94

Figure E.1 Balanced Latin Squares …………………………………………. 126

Figure E.2 Questionnaire Email Screen …………………………………….. 127

Figure E.3 Layout of Questionnaire View ………………………………….. 128

Figure E.4 Module Starter View…………………………………………….. 128

Figure E.5 Sample Question ………………………………………………... 129

Figure E.6 Question-Answer Identification on CA + Bubble ……………... 130

Figure E.7 CA + Bubble Questionnaire UI ………………………………… 131

Figure E.8 Questions on CA + Bubble …………………………………….. 132

Figure E.9 Question-Answer Identification on VSUP + Bubble …………... 133

Figure E.10 VSUP + Bubble Questionnaire UI ……………………………... 134

Figure E.11 Questions on VSUP + Bubble ………………………………….. 135

Figure E.12 Question-Answer Identification on CA + Grid ………………… 136

Figure E.13 CA + Grid Questionnaire UI …………………………………… 137

Figure E.14 Questions on CA + Grid ………………………………………... 138

Figure E.15 Question-Answer Identification on VSUP + Grid ……………... 139

Figure E.16 VSUP + Grid Questionnaire UI ………………………………... 140

Figure E.17 Questions on VSUP + Grid …………………………………….. 141

Figure G.1 Amazon Gift Card Sample …………………………………….. 150

**List of Algorithms**

Algorithm 3.1 MLP Model ………………………………………………..…. 37

Algorithm 3.2 CNN Model ………………………………………….……….. 39

Algorithm 3.3 LSTM Model ………………………………………….………. 41

Algorithm 3.4 ARIMA Model ………………………………………………... 43

Algorithm 3.5 Calculate uncertainty ………………………………………….. 44

Algorithm 3.6 Data scaling …………………………………………………… 44

Algorithm 4.1 CA Construction Formula …………………………………….. 48

Algorithm 4.2 Pattern Generation …………………………………………….. 53

Algorithm 4.3 Texture Generation ……………………………………………… 54