Image classification of radar charts based on SDF framework

Drashti Kanubhai Nayak
Science and Engineering Faculty
Queensland University of Technology
2, George Street, Brisbane, 4001, QLD, Australia

***Abstract*— Image classification is a classification or identification of different measurements of pattern to compare resulting pattern match or mismatch and classify in categories. Image classification is still untouched for various pattern images based on different frameworks of diverse fields of work. In this paper, we are concentrating on how we can introduce pattern recognition in Management System field that follows SDF (Safety Dynamic Framework) with specific radar charts as input data. What we observed based on existing solutions that they are able to classify different types of charts but not detailed data extraction of particular charts based on specific structure. The approach for solution-oriented problem, we needed to evaluate which machine learning algorithm will be more useful for image classification of radar charts based on SDF framework including features and define steps of process required from automated application development to related analysis. Based on multiple existing solutions, it was concluded that CNN (Convolutional Neural Network) is the best existing approach to apply on naïve framework. The resultant accuracy of classification was set to maximum correctly classified instances to number of instances. And, our models is able to achieve at least 90% accuracy in image classification across all 17 categories of SDF.**

Keywords—image classification, pattern recognition, radar charts, SDF(Safety Dynamics Framework), deep learning, management support system

I. INTRODUCTION (*Heading 1*)

Simon Tegg (2014) shows The framework of balanced pattern measurement of strength and weakness of cultural values known as SDF (Safety Dynamics Framework) to measure to address any conflict between targets and receiving result into opinion mining statement. Interestingly, not much is researched or technically done to enhance this framework for company's benefit. This framework is also followed by industry-partner Cultural Radar, a leading safety culture platform. The tasks are already carried out in a semi-computerized system based on radar charts data. However, this is where we identified a business oriented gap and a technical need to research on ways to use pattern recognition algorithm based on SDF framework to address balance and effectiveness of risk and safety factors. Most important change will be a shift from manual to technological operations which saves energy and time with bonus of accurate results.

The objective of this research is to discover a fully automated machine learning algorithm that identifies specifically patterns of radar charts for METES managers as per the given quad framework of cultural dynamics. The concept is new and not examined much before and shows need to enhance management system with technical support. By achieving this, researchers will be able to potentially propose an efficient application of pattern recognition into Management Support System field for further progression on different measurements. Basically, creating a new solution with existing technology is the point of research. However, the new knowledge created by this research will be useful to recognize any pattern based on SDF framework for any management system, even if its application is stopped by Cultural Radar Company. There will be a new study on compared models and result outcome for SDF-oriented data. The process of designing it will also further be able to explain the limitations or future enhancements of the system. At the end, we will have an automated image classifier of radar charts analyzing system based on SDF by examining various models to found most suitable classification model for the required output.

For this project, we found out the pattern recognition system is divided into two parts: 1). Image classification and 2). Data Extraction. For the present, we will focus only on Image classification of radar charts into defined categories by Cultural Radar. The data extraction and prediction of image for new discovered categories will be part of future enhancement. For that, we are exploring existing data models efficient to get insightful image classification as no other researchers have tried to explore radar charts data analysis in management field. Because this solution-oriented problem is exploratory in nature, we will be analyzing different machine learning algorithms and models that will showcase chosen methodology and getting accurate probabilistic image classification solution.

Initially, different machine learning models are examined based on existing systems like ChartSense, WebPlotDigitizer and many others. This will result into giving a clear picture of processes need to be carried out for creating technical application. Later, most efficient model is selected and process of training that model with required accuracy measurement will be carried out. Discussion on the proposed model's implementation steps and processes of machine learning which will be carried out during development phase. Data will be allocated by the Cultural Radar and many data preparation techniques will be carried out to clean data. We are going to use the existing Exception classifier, as coding the model from scratch takes time and examining lot of factors. We might would like to look through other existing classifiers like VGG16, VGG19, Inception, etc. Whereas, we will prefer the most preferred default settings. For safe analyzing, the probability of classes kept to top 5 categories that relates to an image. The current goal is to at least create a solution who is able to classify different images into given 17 classes which further can be extended to predict various categories with

II. RELATED WORK

III. THE MODEL

As described in related work section about the existing pattern recognition systems based chart data types as image. The research is arti-fact methodology where we create a machine learning classifier based on radar charts who can classify images into given 17 categories. And those are, Optimized Equilibrium, Under-Optimized Equilibrium, Over-Optimized Equilibrium, Collective Involvement, Bottom-Up Creative Problem Solving, Top-Down Coordination And Control, Goals And Achievement Focused, Participative Change, Innovative Competitiveness, Collective Compliance, Coordinating Achievement, Tension_1, Tension_2, Where are the Champions, Where are the Agents, Where are the Administrators and Where are the Strategists. These categories help Cultural Radar define the client's safety culture dynamics and how far/near from the required to maintain management. These categories are differently indicated on SDF framework which is a type of radar chart. The SDF is a quad pattern framework where each quadrant depicts a specific theme of cultural dynamics. They are collaborate, create, control and compete. The model will define images of radar charts based on these themes given in figure.1.

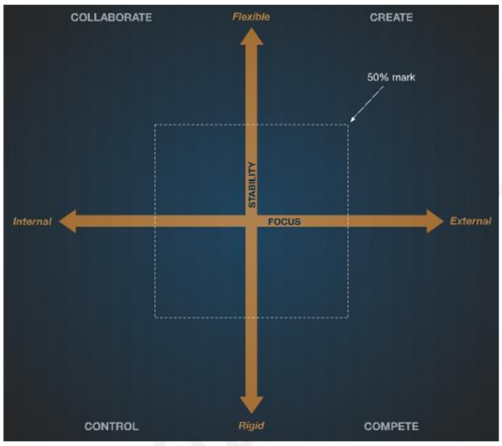

Figure 4 – SDF reference points

The existing solutions and research articles proved CNN to be most accurate model for image classification. CNN is chosen because of its high accuracy for classification images and dimensionality reduction without losing the quality of an image. The CNN model is made up of different layers and the layers can be customized as per needed. The CNN model is consist of input layer, hidden layer consists convolution and pooling layers, one fully connected layer and output classifier layer. The purpose of CNN is to automatically do feature extraction and reduce the dimensionality at pooling layer. At the end, in fully connected layer with the help of softmax activation function hyper parameters tuning to classify images into described classes.

The first part is Feature extraction and later is classification model building. In feature extraction part, various high-level features are learned from input in a sequence of convolution and pooling layers. Convolution layer is the crucial element of CNN architecture. It consists filters that filters raw pixels from extracted features into RGB channel and after computing the dot product gives filtered pixels. We have kept the results into 3-D as in future images the axis of radar charts will be defined in multiple colors' to depict different scenarios of one Client Company. These features later activated in training the model phase. The operation of Convolution layer shapes the network based on size of filters and number of convolution maps. These process increases the capability of CNN to learn different weights and biases of different convolution maps to extract important features. Relu is used as an activation function in this layer which is defined as $f(x) = \max(0, x)$ where x act as an input to a neuron.

The sub layer in convolution layer is max pooling layer. To reduce the dimension GAP (global average pooling) layer used, it also helps in reducing over fitting due to small size of data. It helps in reducing the size of feature map as the output is the maximum value of activation function.

Then comes the fully connected layer which is our classification model. In classification model, the convolution network is created by each neuron fully connected after learning all features from previous hidden layer. The Exception classifier used is considered top most classifier in achieving highest accuracy in image classification. This model was chosen as Cultural radar main focus while achieving was memory reduction and it is one of the smallest weighted models. Xception is specifically a modification of Inception. In Inception, spatial correlations are captured 3x3 or 5x5 convolution, Xception performs 1x1 for every channel and the 3x3 for each output this helps in depth wise separable convolutions

The data is split into default 80% training set and 20 % training set. Epochs are kept till 50 with difference of 5 to understand when the data is fully trained where it observes loss. The default input shape of RGB channel is (299, 299, 3). At the end, layer is fully connected with 1024 units which is activated by relu activation function. The model is training on 258 images and validation on 65 images. In fully connected layer, softmax activation function is used to calculate the scores of classes which gives probability that an image belongs to particular class.

IV. EXPERIMENTAL RESULTS

In this section, we describe dataset used in classifying radar charts and resultant accuracy received for all 17 categories defined on the basis of SDF by management systems. A table is also presented to showcase accuracy for each category achieved.

A. *Experiment Setup*

For implementation of CNN model Tensor Flow (2.3.0) and Keras (2.4.3) libraries are used. The model is built on Acer laptop with Intel i5-5200U CPU @ 2.20GHz and 8GB DDR3 L memory. Laptop runs with Microsoft Windows 10 64-bit operating system. Google Colaboratory free cloud service useful to create machine learning models in python and enabling GPU for faster processing.

B. *The dataset*

For model building, dataset is crucial element in any model building. In CNN, the data is divided in to training set, test set and validation set. Training set contains data which trains model with features selected in data pre-processing phase to get intended output. Test set contains data which is completely new for model to check the predictability of model. And validation set checks entire model accuracy in getting required output.

Due to confidential data of company Cultural Radar, we only received one image per category to define category. The sample size was extremely small for training the model i.e. 17 images for 17 classes. The chart images can't be rotated or flipped due to it changes the meaning of patterns in the data. The description of chart types and categories was given to us by Culture Radar in their requirement report. As above mentioned data was kept in 3D RGB channel to detect different axis colors. The brightness was changed of data. And dataset was increased automatically based on brightness criteria. At the end, dataset was created containing 18 images of different effects per category. The image set was really small hence we couldn't explore different data visualization factors or conceptual diagrams. Hence, we didn't found need to carry out annotation process.

For further, data preprocessing

The resultant accuracy measurement shown in

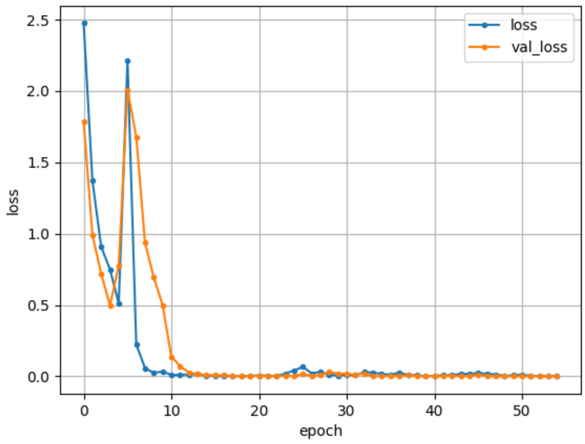

Fig. Loss of accuracy observed from 0 to 50 epochs

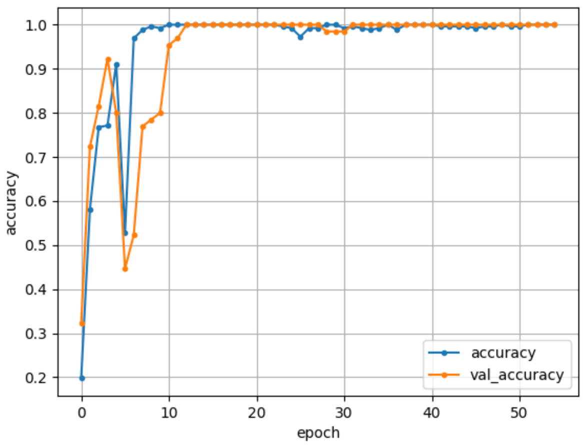

Fig. shows accuracy observed from 0 to 50 epochs.

The above two figure shows the precision measurement of the model. It is seen from the figures that 20 epochs is more than sufficient in training the data. However, when changing the brightness there were few images found around 5-10 iteration where images either being complte black or white showed loss in accuracy and fluctuation is observed in the figure. It hsows that model was confused in showing the relevant probable category.

C. *The result*

Data Augmentation helped in increasing the dataset to build a model look for solution to the problem. Except changing the image factors there were no other differences in images as it doesn't extracted from wide and complex image network. The table is created to show accuracy received per category. More than 90% accuracy was obtained with less feature differences and more than 75% accuracy obtained when image was discolored.

Categories	Accuracy
Optimised Equilibrium	93.56%
Under-Optimised Equilibrium	97.45%
Over-Optimised Equilibrium	91.1%
Collective Involvement	93.75%
Bottom-Up Creative Problem Solving	96.65%
Top-Down Coordination and Control	92.78%
Goals and Achievement Focused	98.20%
Participative Change	92.28%
Innovative Competitiveness	94.45%
Collective Compliance	97.05%
Coordinating Achievement	92.44%
Tension 1	94.52%
Tension 2	93.69%
Where are the Champions	93.52%
Where are the Agents	97.75%
Where are the Administrators	98.15%
Where are the Strategists	96.79%

This accuracy showed probabilities for particular particular image to classified for probable class. However, more than 90% all images were correctly classified for the relevant categories.

CONCLUSION

The focus of this paper is on classification of radar chart images contained SDF framework especially for cultural dynamics used in management system. This research work is motivated by importance of automated classification techniques required in various field and notion that supports technology still needs to reach far-wide areas which are still unexplored. This aids in development of SDF framework that further serves pattern visualization and relevant chart analysis, findings and redesign. The proposed model is based on CNN for radar chart image classification for specific clientele base which classifies 306 images into 17 categories. There were challenges of image resolution, discoloration, variation in sizes, graphics with low variance but preprocessing methods shows efficiency of our model.

It was observe that CNN gave best accuracy for classifying radar chart images that supports our first step to creating fully automated pattern recognition system. However, the limitations of the current model is very small amount of data. The client needs to provide with more different defined pattern images to get more accurate results while solving real-world problems. The data is not only small but also limited to the data provided by company. Even though, we choose Exception classifier based on described feature there are multiple classifiers in the world. A research on exploring different classifiers for the stated data, for stated purpose need to be carried out.

After completion of image classification step, data extraction is next and future enhancement steps which will be fully automated with integrating guided analytics in it. In future enhancements, an automated data extraction and descriptor model must be integrated with the existing proposed model. As that is the second requirement by the Cultural Radar, to get electrically relevant categorical text description of shape, distribution and interpretation of output radar chart image to solve the managerial problems. In that step, we liked to add guided analytics for user interaction; in case needed for exploring undefined data. This bring the enhancement for predicting classification as images will not be restricted to defined 17 categories and might introduce new categories which system needs to predict.

References

[1] Cameron, K. S., & Quinn, R. E. (2011). Diagnosing and changing organizational culture based on the competing values framework (3rd ed.). Jossey-Bass. Page number: 100-125, 221-245.

[2] Kim S. Cameron, Robert E. Quinn, Jeff Degraff, & Anjan V. Thakor. (2014). Competing Values Leadership. Edward Elgar Publishing.

[3]Johannesson, P. (2008). Information Systems Engineering: From Data Analysis to Process Networks, 227-224.

[4] Cameron, K. (2014). Competing values leadership (Second ed., new horizons in management). Cheltenham, UK; Northampton, MA: Edward Elgar

[5] Francois Chollet (2017). Xception: Deep Learning with Depthwise Separable Convolutions. Google Inc.

[6] D. Jung, W. Kim, H. Song, J. Hwang, B. Lee, B. Kim, J. Seo., ChartSense: Interactive Data Extraction from Chart Images. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17). ACM, New York, NY, USA, pp.6706– 6717, 2017

[7]Zara Bandekar (2020). Cultural Radar SDF project requirements and system description report

[8]Simon Tegg (2014). Pattern Recognition for Management. https://blog.loomio.org/2014/01/28/pattern-recognition/

[9] Sven A. Carlsson(2008). Determining Requirements for Management Support Systems. chapter 9, pages 207-228

[10]Huang, Y. M. K. W., Lin, C. C., & Wu, Z. X. (2019). A deep learning and image recognition system for image recognition. Data Sci. Pattern Recognition. 3(2), 1-23.

[11]Anil K. Jain, Robert P.W. Duin, and Jianchang Mao (January 2000). Statistical Pattern Recognition: A Review. IEEE Transactions on Patten Analysis and Machine Intelligence.

[12] Priyanka Sharma and Manavjeet Kaur (Volume 3, Issue 4, April, 2013). Classification in Pattern Recognition: A Review. International Journal of Advanced Research in Computer Science and Software Engineering, India.

[13] Marc Velay and Fabrice Daniel (June 2018). Stock Chart Pattern recognition with Deep Learning. Artificial Intelligence Department of Lusis, Paris, France.

[14] Dan Kaczynski, Leigh Wood and Ansie Harding. Using radar charts with qualitative evaluation. SAGE publication.

[15] Mahmoud A. Barghash and Nader S. Santarisi (2004). Pattern recognition of control charts using artificial neural networks – analysing the effect of the training parameters. Kluwer Academic Publishers, Netherlands.

[16] Carlie Idoine, Peter Krensky, Erick Brethenoux, Jim Hare, Svetlana Sicular, Shubhangi Vashisth (February 2018). Magic Quadrant for Data Science and Machine-Learning Platforms. Gartner.

[17] R. Begg, J. Kameruzzaman (2004). A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data.

[18] W. Huang, C. L. Tan, W. K. Leow, Model-based Chart Image Recognition, GREC, 2003

[19] A. Rohatgi. 2018. WebPlotDigitizer, Version 4.1. https://automeris.io/WebPlotDigitizer/

[20]M. Savva, N. Kong, A. Chhajta, L. Fei-Fei, M. Agrawala, J. Heer., ReVision: Automated classification, analysis and redesign of chart images. In Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11), pp.393-402., 2011

[21] M. Cliche, D. Rosenberg, D. Madeka, C. Yee. 2017. Scatteract: Automated extraction of data from scatter plots, unpublished.

[22] Keras. 2018. | Keras: The Python Deep Learning library https://keras.io/

[23] Tensorflow. 2018 | An open source machine learning framework for everyone https://www.tensorflow.org/