Webology (ISSN: 1735-188X)
Volume 18, Number 6, 2021

A Ml Approach Of Capturing, Augmenting And Utilizing Online Trends To Improve Textile Designs Patterns

S.Senthilvel[1] , Dr. A.Prema[2]

[1]Research Scholar, Department of Computer Science, Vels Institute of Science Technology & Advance Studies, Chennai.

[2]Associate Professor, Department of Computer Science, School of Computing Science, Vels Institute of Science Technology & Advance Studies, Chennai.

Abstract

Designing competitive design patterns in textile industry to improve sales is always a difficult creative task, where the designer knowledge need to be up-to-date with the latest trends and peoples preferences. This paper attempts a novel method of generating trendy design patterns with the help of supervised machine learning algorithms. A live dataset of ongoing trends will be collected from twitter hashtags through a customized tweepy API. The popularity of the images shared along with the hashtags are categorized through Content Based Image Retrieval (CBIR) technique. Various image augmentation techniques are used to generate different combinations of designs from the categorized pool of the popular images. Finally a Deep Neural architecture is developed to predict the best design patterns through supervised training. After implementation, a comparative analysis of various optimizations in neural architectures for developing popular patterns will be studied.

INTRODUCTION

Social medias are the reflecting mirror of peoples trends in which terabytes of information are generated, shared and consumed by millions of people in daily basis. Many bots, crawlers and scraping tools are available in market to capture the online trends and convert it into meaningful data. Many social media platforms are developed their own application interface(API) and made them available to public withcertain terms and conditions. This help researchers from all over the world to legally acquire the regulated data from social media. All around the world research works are published for utilizing the social media trends to improvise various aspects of technological innovations.

The emergence of cutting edge machine learning algorithms plays a vital role in shaping the next generation world with its many applications in medicine, physics, transportation, space research etc. There are vast number of fields still can be improvised with the integration of ML algorithms with its application. To design a successful ML model it requires huge amount of training data set. So by

embedding huge datas accumulated from social media and data hungry Machine Learning algorithms will results in AI revolution. Yet machine learning models has a limitation of lacking self-inspiration and creativity, it can work with existing model as training reference to predict the target with specific parameters. But in the case of drafting a trendy design in textiles, existing models will be useless because of the constant changes in social trends & consumer taste. Improving sales in textiles requires enormous unique and trendy design patterns for which many creative designer minds need to be employed.

I. Related Works

Social media are privileged mediums for generating rich and multidimensional big data's which exactly reflects an average user mood, requirements and priorities. It provides an unprecedented multifaceted insights to the firms and stakeholders to understand the commercial trend and leads to client-centric innovations. Social media acts as a bridge between the people and companies for co-creating ideas and concepts and supporting new product launches [1].

A literature study of finding the impact of social media platforms in the research field is explained in detail. By using Google scholar as source, various research scholars publications relevant to the social media trends and platforms are gathered and categorized during the publication period of 2007-2013.Over 229 articles are found to be published over the span of 6 years which emphasis the dominance the social media trends in the field of big data and research works [2].

Web scraping is a technique of using internet tools to extract the unstructured data from a website and to convert in to a meaningful structured data. Using DOM parsing, web scraping tools, html tags inside a website will be converted into an JavaScript object notion (JSON) or Extended Markup Language (XML) data object. Which can be interfaced with any programming language. This scraping [3] allows the programmers to analyse various aspects of the website ex: Change detection in website, Product price comparison on multiple sites, Web indexing & Rank checking etc…

The polarity of the user trends about a particular issue can be categorized using sentiment analysis. Tweepy [4] Application Interface (API) acts as a bridge between the twitter social media platform and the sentiment analysis python algorithm. Tweepy API is used to find and extract the top trending hashtags from social media and predict the polarity by analysing the tone of the text using Natural Language Processing Toolkit (NLTK).

A comprehensive review on selenium automation web testing tool is presented [5]. Selenium is used for performing numerous testing analysis and to reduce the manual testing time. Selenium can automate the functional and unit testing and it can provide a detailed error and response timing report.

Image retrieval is a technique of finding semantic feature from an image and searching the database of other images which correlates with the image features. A better matching accuracy can be achieved [6] using Gaussian filter at the pre-processing stage to remove speckle and salt pepper noises. A retrieval

rate of 94.90% is achieved by employing multi SVM classifier in PNN architecture. Color features are extracted using color co-occurrence method and grey features are extracted using grey level co-occurrence (GLCM) method.

Machine learning problems can be categorized into three major types: classification, clustering and regression. The type of machine learning methodology will be selected based on the nature of the problem and the type of the available training set data [7].

Convolutional neural networks can be trained to identify the meaning full patterns. Better image/pattern classification is achieved by developing convolutional patches [8] for identifying particular visual pattern. Using a fully connected neural network better image classification is achieved to identify similar image patterns which will be helpful in textile designing to spot the repeated texture designs and to rectify them.

Due to the rapid growth in Textile E-shopping a knowledge database of categorizing the user's like and don't likes can be indexed, the preferred design and textures are periodically identified and it influences the trending fashion developments.Style2Vec architecture is proposed [9] to develop entirely new design pattern from the highlighted huge assortments of client made style sets.Style3Vec uses two different CNN systems to perform likelihood pattern arrangements with reference to client storage room preference for different semantics and dimension.

Several challenges and concern need to be addressed while designing a neural architecture, based on the nature and dimensionality of the input vectors, the total number of input nodes, number of hidden layers, neurons per layer, type of activation and loss function, optimization algorithm, finally total number of epochs and processing batch size for each epoch need to be determined [10]. Tweaking any of the above parameters will results in different prediction accuracy, some factors might have less impact on the accuracy, some leads to huge prediction changes.

II. Design Flow

Simplified steps of the design flow is illustrated in the Fig.1. Tweepy API for python can pass input parameters to the Twitter web portal and can extract the relevant popular hashtags based on specific location supplied in parameters. Dynamic page contents are observed in live using selenium testing tool, the number of valid images/hashtag and its downloadable urls are scraped through beautiful soup python library.

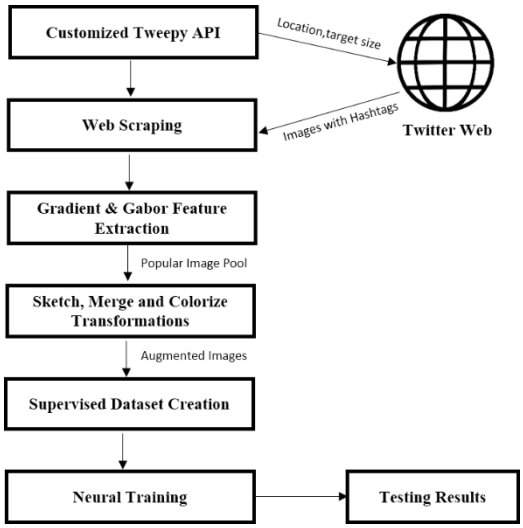

Fig.1 Design Architecture

III. Trending Image Capturing & Improvisation

Initially a twitter account with developer privileges is created, by answering a series of developer account related questions, a request of twitter external authentication Key will be created from the developer end. Using the twitter Auth Handler and Access Token any twitter account feed can be accessed programmatically through tweepy API. Using trends place function in tweepy , we can find and extract the trending information of a particular area, ID denotes the where on Earth ID(WOEID). A **WOEID** (Where On Earth Identifier) is a unique 32-bit reference identifier, originally defined by Geo Planet and now assigned by Yahoo [],which is used to extract the trending hashtags in twitter. After finding the url link of each user status, using Beautiful soup python web scraping package the status wall is combed for any image postings under that particular trend. The resultant image links will be stored as a List. Each image in the list will be extracted inside the folder named the hashtag.

a. Popularity Categorization using CBIR

The scavenged images inside the Hash tag will act as database repository, the goal is to find top n images trended repeatedly, to achieve that Content/Feature based image matching will be done. Hog feature is used to match the objects present in the query image and compare the objects in the database images, the number of matches will be stored inside a python dictionary. The histogram of oriented gradients (HOG) is a feature descriptor used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image

$$g = \sqrt{g_x^2 + g_y^2}$$
$$\theta = \arctan \frac{g_y}{g_x} \quad \textbf{(1)}$$

In pre-processing the images are resized to 100x100 pixel to avoid timing complexity this can be

avoided in case of systems running with GPU's . The pre-processed image is then subjected to HOG calculation and feature descriptors are referred as 'fd'. The multichannel parameter must be True for color images and False for grey scale images. Feature descriptor values of all the other images are referred as 'cfd', the correlation confection between feature vector of the query image(fd) and current search image (cfd)will be compared

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad \textbf{(2)}$$

Where x_i denoted column pixel and y_i row pixel

b. Augmented Transformation

An array of design synthesizer which includes transforming the input image into a water color painting, a stylized pencil sketch, altering the color map values to change the radiance and merging the image into predefined patterns.

Fig.2 a) Original image b) Gray pencil sketch c) colored pencil sketch

These augmentation images are generated with arbitrary parameters like sigma strength, shade factor, color map array etc... So no image will be the same type of the previous.
These augmented images forms the untrained datasets, which will be further subjected to supervised dataset, presently the best augmented images are picked by offline users, in future to get a global appeal, these most trended and augmented imaged can be available online for user rating, which will greatly improve the training datasets.

IV. Neural Training & Testing

Neural networks earned a great reputation in the field of supervised learning because of its versatility for adapting to different input criteria's. In our work, It requires a labelled dataset of augmented image features with what the user likes and dislikes. This dataset is collected with a help of TKinter python GUI library.

The randomly generated design patterns are subjected to user review with a like/discard button, this builds the labelled datasets for supervised neural training.

Training Model & Testing

A Neural architecture is designed with python keras library with tensor flow as backend, Tensor flow is a open source AI development platform developed by Google as a Software as a Service (SAAS). Tensor flow can handle big data with the help of map reduce parallel processing, it also supports map reduce as a backend function.

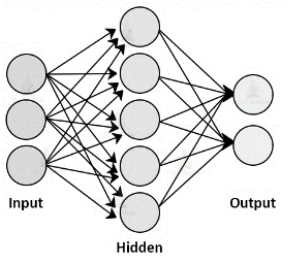

Fig.5. Fully Connected Neural Architecture

Initially conventional neural architecture of single hidden layer with 4 nodes per layer is configured for training and testing. To reduce the timing complexity the once trained bias and weights values of all layers are stored as a separate Hierarchical Data Format (H5) File. The training and testing is done with different levels of hidden layers and also with different node size. To find the best architecture for this user guided dataset which various from user to user, the loss and optimizer functions are calibrated and the resultant accuracy level has been tabulated.

Nodes/Hidden Layer	1 Hidden Layer(%)	2 Hidden Layer(%)	3 Hidden Layer(%)
2 Nodes/Layer	61	82	88.4
4 Nodes/Layer	63	83	88.7
8 Nodes/Layer	71	84	89.1
16 Nodes/Layer	74	85.5	89.7
32 Nodes/Layer	76	86	90
64 Nodes/Layer	79	87	91
128 Nodes/Layer	82	88.3	92

Tab 1. Accuracy over Nodes and layer

Conclusion

A Novel idea of extracting and utilizing trendy social media images as textile print design is successfully implemented with the help of various open source languages and tools. Twitter is preferred as experimental platform because of its elegant and simple design, which has good spam

filter and fake id identification to avoid promoted/fake trends from a single user id. The trending hashtags are identified by importing tweepy API into python script. Selenium acts as a browser agent to auto navigate to the each hashtag and collecting the image associated with it.

A 97% better image categorization accuracy is achieved in Content based Image retrieval through HOG feature extraction method. Since it follows a simple correlation of maximum shared images, pre-processing and complex classification mechanism are not needed which greatly reduces the searching time. The classified images are subjected to random image augmentation comprised of stroke, color map and pattern merging methods. A neural guided pattern merging algorithm is used to achieve smooth blend of the image with the background design patterns,

Because of its novelty to perform all the calculation with live data sets, there is no benchmark dataset is available to compare the accuracy with previous methods. So various neural architectures are tried in trial and errorto find the best architecture to give maximum accuracy based on user likes. Adam optimizer is used to calculate the weights and learning late if each neural node. Rectified Linear unit (Re Lu) is used as output function to covert the sigmoid decimal output into binary (yes/no).A maximum epoch time for 200 iterations are configure for training and optimized results are achieved in 152 epoch itself. Increasing more than 3 Hidden Layer doesn't yield much difference in prediction accuracy but increases the prediction delay. Thus a successful design and implementation of utilizing ML algorithms and Automation tools for design user centric textile design is done and tested.

It opens a vista of future possibilities, In which the training of the proposed neural networks can be made public as a Software as online service , so millions of design ideas can be pooled from all around the world to develop a AI master in fashion designing.

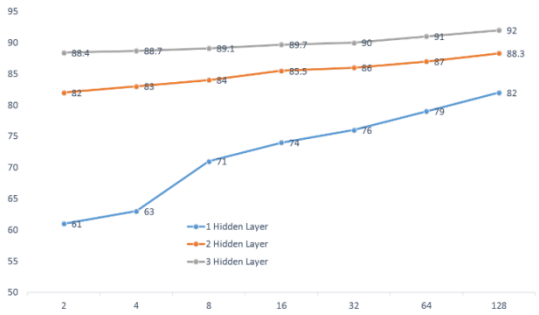

Fig 6. Comparison of Neural Architecture

References

[1] Bhimani, H., "Social media and innovation: A systematic literature review and future research directions", Technological Forecasting & Social Change 2018.

[2] Chareen L. Snelson.,"Qualitative and Mixed Methods Social Media Research: A Review of the Literature", International Journal of Qualitative Methods January-December 2016.

[3]Anand V. Saurkar.,"An Overview On Web Scraping Techniques And Tools", International Journal on Future Revolution in Computer Science & Communication Engineering 2018,Volume 4,Issue 4.

[4]Rahul Pandya., "Polarity Testing and Analysis of tweets in Twitter using Tweepy" ,International

Journal of Computational Engineering Research 2021, Volume 11,Issue 2.
[5]Deepthi Wilson. R., "A Comprehensive Review on Selenium Automation Testing Tool", International Journal of Engineering Research & Technology 2017.
[6] Mudhafar J. J. "Content-Based Image Retrieval of Color, Shape, and Texture by Using Novel multi-SVM Classifier", International Journal of Machine Learning and Computing 2019, Volume 9.
[7] Susmita Ray., "A Quick Review of Machine Learning Algorithms", International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (Com-IT-Con), India, 14th -16th Feb 2019,pp.36-39.
[8]Hongzhi Li., Joseph G. Ellis , Lei Zhang and Shih-Fu Chang, "Pattern Net: Visual Pattern Mining with Deep Neural Network", ICMR ’18,2018.
[9] Hanbit Lee., Jinseok Seol, Sang-goo Lee,"Style2Vec: Representation Learning for Fashion Items from Style Sets", Computer Vision and Pattern Recognition 2018.
[10] Ochin Sharma, "Deep Challenges Associated with Deep Learning", International Conference on Machine Learning, Big Data, Cloud and Parallel Computing 2019.