Brain-oriented Cconvolutional Neural Network Computer Style Recognition of Classical Chinese Poetry

Jing Xuan[1], Zhongshi He[1*], Liangyan Li[2], Weidong He[3], Fei Guo[1], Hang Zhang[1], Qiong Wu[4]

ABSTRACT

This paper aims to develop a feasible way to recognize the style of classical Chinese poetry with computers. To this end, the authors explored the connectionism in neuroscience, and explained the cognitive word embedding with the convolutional neural network (CNN). On the one hand, the genetic algorithm was adopted to extract keywords from traditional hand-labelled and selected information; on the other hand, a novel computer learning method was proposed based on text-to-image (T2I) CNN for big data. The proposed method was contrasted with the traditional genetic algorithm of naive Bayes and information gain. The experimental results show that our method achieved better classification accuracy with less human intervention than the traditional genetic algorithm. Hence, the CNN-based method is feasible on big data, both in theory and practice. This cross-disciplinary practice sheds light on stylistics, literature engineering, poetry cognition and neural network projects.

Key Words: Convolutional Neural Network (CNN), Classical Chinese Poetry, Cognitive Engineering, Connectionism

DOI Number: 10.14704/nq.2018.16.4.1214 **NeuroQuantology 2018; 16(4):107-115**

Introduction

Convolutional neural network

Cognitive engineering is a method for applying cognitive psychology and relevant disciplines into the actual operation of human-computer systems. It has been extensively implemented in linguistics, neuroscience, artificial intelligence, etc. However, the big data of poetry, especially classical Chinese poetry, has posed a great challenge to cognitive engineering. The existing strategies for natural language processing (NLP) lacks sufficient ability to process raw data, and requires carefulness and expertise on feature extraction.

With the advent of deep learning, many unsupervised or semi-supervised learning methods have been developed. These methods support the automatic raw data processing (e.g. feature detection and text classification) with computers, thus eliminating the need for human intervention in rule-based representations of the NLP. Hence, deep learning has offered a feasible alternative to handle the big data of poetry with much less emphasis on human ingenuity and prior knowledge. Taking the backpropagation neural network (BNN) for example, the computers automatically learn about the intricate structure of large datasets using the backpropagation algorithm, while updating the internal parameters from layer to layer. Similarly, the reverse image search of Google combines the word embedding and image representation into a

Corresponding author: Zhongshi He
Address: [1]College of Computer Science, Chongqing University, Chongqing 400044, China; [2]Research Center for Language, Cognition and Language Application, Chongqing University, Chongqing 400044, China; [3]College of Computer Science, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; [4]College of Mechanical Engineering, Chongqing University, Chongqing 400044, China
e-mail ✉ zshe@cqu.edu.cn
Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
Received: 7 February 2018; **Accepted:** 4 April 2018

search tool for associated text and images (Weston, Bengio, & Usunier, 2010).

The creation of Boltzmann machine (Srivastava and Salakhutdinov, 2012) and deep belief network (Hinton, 2009) has brought deep-learning neural networks, recurrent or convolutional (Bengio *et al.*, 2013), to the spotlight. The recurrent neural network (RNN) focuses on sequential data like text and speech, while the convolutional neural network (CNN) excels in processing images, audios and videos (Lecun *et al.*, 2015). Specifically, the CNN refers to a type of feedforward network that processes data in multiple arrays by backpropagation algorithm. The network consists of convolutional layers and pooling layers. The former captures feature maps and passes local weighted sum to the subsequent layer, while the latter reduces network dimensions through merging of similar features.

With a specialty in image processing and face recognition, the CNN has been applied to file reading, optical character recognition and handwriting recognition since the 1990s. In practice, the network avoids overfitting and achieves stable performance with its large datasets and powerful models. Compared to the RNN, the CNN faces little difficulty in training and consumes a rather short period of time. This is because of the limited number of connections and parameters in the classification tasks of the CNN. Inspired by neuroscience, the latest CNN was founded to respond to visual images as efficiently as monkey's inferior temporal (IT) cortex (Cadieu *et al.*,2014). The proliferation of GPUs makes it possible to train large labelled datasets through unsupervised learning with an optimized 2D CNN, e.g. the ImageNet (Krizhevsky *et al.*,2012) in the ILSVRC-2010 and ILSVRC-2012 competitions.

Classical Chinese poetry

The classical Chinese poetry, dated back to 2,500 years ago, has an immense influence on poetry worldwide ("classical Chinese poetry" n.d.). The stylistic development of classical Chinese poetry is often divided into several eras in chronological order, namely the early (pre-Qin dynasty) classical Chinese poetry (e.g. the Chu Ci and the Shi Jing), Yuefu (Music Bureau) poetry, Six Dynasties poetry, Sui poetry, Tang poetry, Song Ci-poetry, Yuan poetry, Ming poetry, Qing poetry, and post-Qing poetry.

Among them, Tang poetry and Song Ci-poetry are the two peaks of Classical Chinese Poetry. The Tang poetry is noted for its rhymed and parallel structures. The largest collection of Tang poetry is Quan Tangshi (Complete Tang Poems), which contains some 49,000 lyric poems by more than 2200 poets (Qian and Huang, 2015). The Song Ci-poetry is a kind of lyric poetry using a poetic meter based upon certain patterns of fixed-rhythm formal types, of which there were at least 1000 of these set patterns, each associated with a particular title. Originally, Ci was written to be sung to a specific tune of that title, with set rhythm, rhyme, and tempo ("classical Chinese poetry" n.d.). The Ci poems are generally divided into two styles: bold and unconstrainted (Haofang) and graceful and restrained (Wanyue) (Li, 2004). Below are two examples of the same title "Snow Tune" for the two styles:

Haofang style:
Snow Tune: "Spring in a Pleasure Garden" (Mao Zedong)
See what the northern countries show: Hundreds of leagues ice-bound go; Thousands of leagues flies snow.
（北国风光, 千里冰封, 万里雪飘 。）
Wanyue style:
Snow Tune: "Snow" (Li Liangyan)
Behold the view of a little town: one league of frozen ice, and two leagues of drifting snow. (L. Li, He, & Yi, 2005)
（小城风光, 一里冰封, 两里雪飘 。）

Considering the massive amount of classical Chinese poems, it is impossible to exhaust all the character combinations in one corpus by the traditional genetic algorithm. For example, the EMNLP 2014 alone collected over 280,000 classical Chinese poems, which contained 18 million Chinese characters (Zhang and Lapata, 2014).

Research problem

Facing the limited amount of labelled characters, transfer learning with parameter fine-tuning was adopted to draw merits from the abundant data of classical Chinese poetry (Ge and Yu, 2017). It is assumed that (1) all classical Chinese poems share the same probability distribution of vector features and hidden constructive rules and (2) all these rules can be captured or learned by computers, using the cognitive engineering method based on the connectionism in neuroscience. The learning process aims to produce a large input training dataset, while minimizing the need for manual labelling. More samples were obtained by adding noises to training, ranking the data by weak classifiers (Parkhi *et al*., 2015) or transforming the images. Then, the reverse pipeline approach (Bird *et al*., 2009) was employed to convert concepts into speech, with the aim of raising computing creativity (Corneli *et al*., 2015). The reverse maximum matching, a common tool for segmenting Chinese words, was also adopted here.

To develop a feasible way to recognize the style of Classical Chinese Poetry, this paper probes deep into the connectionism in neuroscience, and explains cognitive word embedding with the CNN. On the one hand, the genetic algorithm was adopted to extract keywords from traditional hand-labelled and selected information; on the other hand, a novel computer learning method was proposed based on Text-to-Image (T2I) CNN for big data. This research makes the following contributions: First, a new approach was put forward to transform Classical Chinese Poems into images, and classify the poems into Haofang and Wanyue styles with fewer human interventions; Second, the character order of each training poem was reversed to double the size of training dataset in T2I CNN learning; Third, the computers were given the cognitive power to unearth hidden poetry features without interaction of expert rules, like AlphaGo Zero (Silver *et al*., 2017). This research sheds new light on stylistics, literature engineering, cognitive poetry and neural network projects.

Literature Review

In computational linguistics, poems are often split into 3 main levels (Delmonte, 2016): phonetic relational level, poetic relational level and syntactic-semantic relational level. Over the years, numerous parsers have been developed for poetry analysis, including but not limited to information dependency model on syntax parameters (Li and He, 2007), and an automatic analyser of poetic structure and rhythm for Shakespeare's sonnets. The semi-supervised learning methods are also applied in poetry analysis. For instance, Rahgozar and Inkpen (2016) classified Hafez's poems in chronological order with Support Vector Machine (SVM) and Latent Dirichlet Allocation (LDA). The evolution of the other poem processing techniques is reviewed below.

In 1986, Hinton introduced the distributed representation of concepts (Bengio *et al*., 2013), in light of the success of representation learning in object recognition, signal processing and image classification. Then, Hinton and Paccanaro (2001) proposed linear relational embedding after using gradient ascent to maximize a discriminative goodness function. Bengio *et al*., (2003) invented statistical language modelling and applied it in neural network language models (2008). Collobert *et al*., (2011) created a convolutional structured system named SENNA; the system learns internal representations on large unlabelled training data, and targets at general language modelling tasks like syntactic parsing, semantic role labelling, part of speech tagging, and named entity recognition.

In China, Liu and Luo (2016) tracked Tang and Song poems in a biographical database on social networks among poets. Qian and Huang (2015) acquired topic categories and theme evolution through pointwise mutual information and LDA experiments. Hu and Zhu (2015) classified Tang poems into different themes with naive Bayes classifier and SVM algorithms. In addition to the SVM (Li *et al*., 2005; Selangor 2012), information gain (Yi *et al*., 2007) and genetic algorithms (Manurung *et al*., 2009; Zhou *et al*., 2010) have been widely used for computer classification and generation of literature, especially for Classical Chinese Poems.

With the improvement of GPU performance, the neural network engineering has been introduced to the analysis of Classical Chinese Poems (Zhang and Lapata, 2014). For example, the big data-based RNN structure was implemented to import poem generation rules to computers, and achieve raw input without human labelling. Special concerns have been given to the ethics and feasibility of computer poetry writing, especially the tones and rhymes in Jueju (a poem of four lines), Lvshi (a poem of eight lines with a strict tonal pattern and rhyme scheme) and Ci-poetry (Lin, 2013). The research hotspots include the framework, intention, emotion, terms, imageries, bibliography, genre, aesthetic beauty, political stand and geographical migration (Liu, 2013; Wang, 2017; Liu and Luo, 2016; Qian and Huang, 2015; Sundararajan, 2004).

To sum up, the existing efforts on Chinese poem classification or generation are grounded on the RNN. The time-efficiency and network performance can be improved by replacing the RNN with the CNN. For instance, Li *et al*., (2017) generated poem titles rapidly based on CNN semantic relevance. Since the CNN boasts a more stable performance of image processing (Parkhi *et al*., 2015), the classification speed will be accelerated without sacrificing accuracy once the Chinese texts are converted into images (Cai *et al*., 2016).

Connectionism is a movement in cognitive science that hopes to explain intellectual abilities using artificial neural networks. The central connectionist principle is that mental phenomena can be described by interconnected networks of simple and often uniform units. The form of the connections and the units can vary from model to model. For example, units in the network could represent neurons and the connections could represent synapses, like in the brain of a human being. The experiments on these models have demonstrated an ability to learn such skills as face recognition, reading, and the detection of simple grammatical structure ("Connectionism," 2015).

The idea of connectionism has been applied to the analysis of classical Chinese poems. Liangyan Li proposed a term connection model for poetry style recognition and implemented the model in a syntax tagging system (Li, 2009). Delmonte discussed the ways to process styles in the Computing Poetry Style (2013). Yi (2005) tested the computer-aided couplet writing in Chinese poem composition. Through the literature review, the authors attempted to apply the latest methods of neural network engineering onto classical Chinese poetry classification, aiming to achieve fast and accurate computer learning and style recognition.

CNN-Based Poetry Style Recognition

Artificial Neural Network

Neurolinguistics ("Neurolinguistics," n. d.) is the study on the neural mechanism in the human brain (figure 1) that controls the comprehension, generation, and acquisition of languages. The central idea is that all language theories must correspond with the facts of brain neuroscience. In other words, the language models should abstract the brain functions and nerve cell structures in a precise manner.

Over the years, the neuro-linguists have studied the physiological mechanisms by which the brain processes information related to language, and evaluate linguistic and psycholinguistic theories, using aphasiology, brain imaging, electrophysiology, and computer modelling ("Neurolinguistics," n.d.). In human brain, the key to linguistic functions (e.g. language generation and comprehension) lies in neurons, which transmit information via electrical signals.

The neurons are connected in different patterns with constantly changing strength, thus forming different brain functions, such as vision, phonation, and the like. The connection between neurons empowers huge potentials. In particular, the connectionism offers an alternative to the traditional understanding of the brain: that the brain is akin to a computer processing a symbolic language.

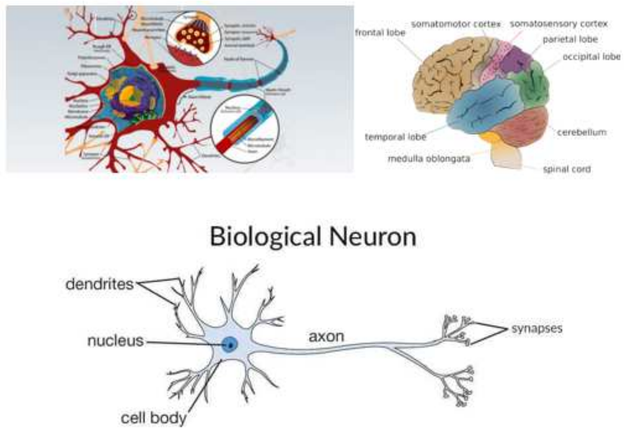

Figure 1. Neurons in human brain
(Source: http://www.humanbrainfacts.org/neurons-in-the-brain.php)

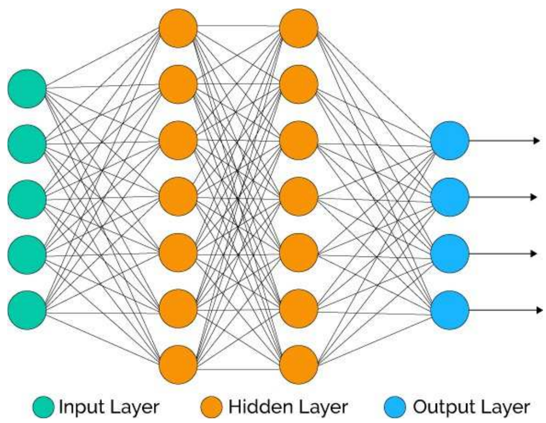

Figure 2. Neurons in an artificial neural network
(Source: https://medium.com/@xenonstack/overview-of-artificial-neural-networks-and-its-applications-2525c1addff7)

In cognitive engineering, the nodes of a neural network are divided into the input layer, output layer and hidden layer (Figure 2). To simulate the human brain, the input layer nodes should represent the sensory neurons, the output layer nodes should represent the motor neurons, and the hidden layer nodes should represent all the other neurons. The network structure should be activated by a certain stimulus, and pass parameters forward or backward through activation functions.

T2I CNN

To give computers the ability of recognizing poetry styles, the poem texts must be converted into image pixel matrix and be input into a CNN structure. For this purpose, the Word2vec was adopted as the encoding method. By this method, each Chinese character was expressed as a unique decimal vector with 200 dimensions in the entire Classical Chinese Poem Database. Each piece of poem was converted into a pixel matrix by adding up the decimal vectors of characters line by line. The resultant poem matrix was set to the size of 260*200 (Table 1). After the conversion, each Chinese character occupied a line of 200 dimensions coded by Word2vec. If a poem exceeded 260 characters, it was split into two 260*200 matrices sharing the same poem label. If a poem fell short of 260 characters, the empty dimensions were filled with zeros. Unlike greyscale images (in which the pixel values are usually integers), the Word2vec-coded values in the poem matrices are decimals. The values reflect the implicit relations between characters in the poem, similar to neuron connections in human brain. These hidden connections were fully extracted by the CNN structure. After the text-to-image transformation, a CNN structure (Figure 3) was designed for the poetry style recognition of classical Chinese poetry.

The next step was to provide abundant training data for CNN learning. The problem was that the number of recorded classical Chinese poems takes up only a fraction of all Chinese poems. To find more training data, the character order of each training poem was reversed to double the size of training dataset. In this way, a piece of poem was entered into the CNN structure in both forward and backward directions, forming a bi-directional CNN input. This method is a common practice in bi-directional RNN structures for Natural Language Processing (NLP). In our poetry style recognition, the bi-directional input does not break the connection between characters, making it possible to classify the poems accurately based on character connection frequency and obey CNN image conversion rules. The feasibility of the method is to be verified in the subsequent experiment.

Table 1. T2I pixel matrix of *On the Stork Tower* (《登鹳雀楼》)

白：0.1234 0.2564 0.2257 … 0.5387
日：0.2587 0.2558 0.2525 … 0.2258
依：0.2622 0.4684 1.2597 … 0.2576
… … …
一：1.7767 0.5767 0.2182 … 0.2258
层：0.2528 0.2598 0.2002 … 0.2568
楼：1.2577 0.2682 0.5892 … 0.2569

登鹳雀楼 王之涣	**On the Stork Tower** *Translated by* Yuanchong Xu
百日依山尽，	The sun along the mountain bows;
黄河入海流。	The Yellow River seaward flows.
欲穷千里目，	You will enjoy a grander sight,
更上一层楼。	If you climb to a greater height.

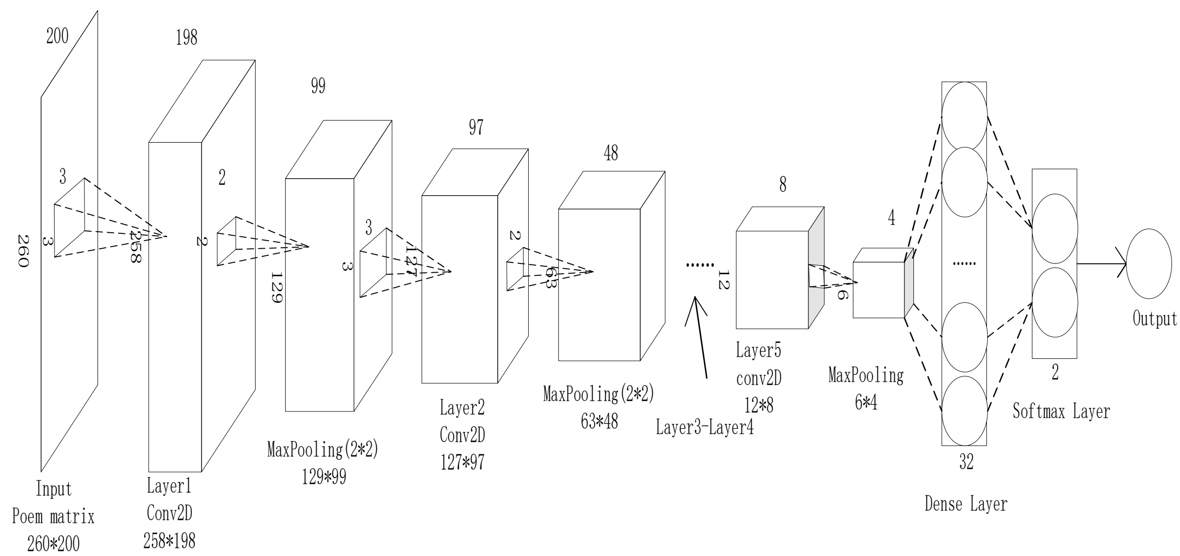

Figure 3. T2I CNN structure

The proposed T2I CNN structure has a total of eight layers.

Layer 1. 128 feature matrices were extracted from 128 convolution kernels by the activation function Relu, and the data amount was reduced by MaxPooling without changing the features.

Layer 2. 64 convolution kernels were used to extract features by Relu, and the redundant data were removed by MaxPooling.

Layer 3. The same with Layer 2. with convolution kernels 64.

Layer 4. The same with Layer 2, except that the number of convolution kernels was changed to 32.

Layer 5. The same with Layer 2, except that the number of convolution kernels was changed to 16.

Layer 6. The previous data were reduced into a 1D matrix.

Layer 7. The network depth was increased by DenseLayer into a 16*1 matrix.

Layer 8: The data were treated with a quadratic classifier using Softmax.

Each input data x was convoluted with a 3*3 convolution kernel:

$$x_0=w*x_i+b \quad (1)$$

where x_0 is the output of each layer; x_i is the input of each layer; w is the weight of convolution kernel; b is the bias. The output nonlinearity was increased by Relu:

$$x_0'=\max(0, x_0) \quad (2)$$

Then, MaxPooling was adopted to extract features from poetry matrix. Next, the data were compressed to 1D by FlattenLayer, and classified into target categories by DenseLayer and Softmax.

Methods

Dataset

For the sake of universality, the authors collected as many classical Chinese poems as possible. The entire corpus consists of EMNLP 2014 poem collection and our manual supplement, adding up to a total of 18,534,445 Chinese characters in 284,899 pieces of poems. The poems range from the pre-Qin dynasty era all the way to modern China. For Song Ci-poetry, 2,509 Haofang poems and 3,091 Wanyue poems were obtained through expert labelling, extracted from books on classical poetry, and downloaded from well-known poetry websites. The experts invited for labelling are either masters or professors of literature. In total, there are 5,600 pieces of poems labelled Haofang or Wanyue. They were divided into training set, validation set and test set.

Genetic algorithm on information gain

199 pieces of Haofang poems and 210 pieces of Wanyue poems were selected from the Complete Collection of Song-Ci Poems. Then, the characters were extracted by experts from these poems together with character frequency and co-occurrence. After that, the classification features were learned by naive Bayes, and poem style was reflected in character features. Experimental Results in figure 4 showed a relatively stable result when feature number reaching 100 or so. The information gain was applied to evaluate correlations between each character and its target style category. The most closely-related 100 characters were obtained for the genetic algorithm to determine the poetry style of each input poem. These characters include:

“花”，“山”，“春”，“何”，“头”，“生”，“中”，“万”，“去”，“月”，“霜”，“雨”，“愁”，“然”，“节”，“塞”，“地”，“西”，“将”，“江”，“车”，“四”，“北”，“方”，“堂”，“无”，“悲”，“梦”，“台”，“人”，“兰”，“渚”，“数”，“争”，“兴”，“声”，“别”，“胡”，“扬”，“淮”，“可”，“思”，“吹”，“酒”，“坐”，“故”，“出”，“碧”，“雷”，“乡”，“香”，“会”，“曲”，“看”，
… … …

These characters were embedded in the genetic algorithm through the following steps:

```
{
gen = 0;
initialize();
evaluate(P(gen));
do {
gen = gen+1;
generate new population P(gen) from P(gen-1)
through select, crossover,
and mutation;
evaluate(P(gen));
} while (gen <= maxgen)
}
```

The experiment shows that the classification accuracy was stabilised when the feature selection involved around 100 characters.

Based on naive Bayes, the poetry style classifier was trained by 10-fold cross validation of 199 Haofang and 210 Wanyue poems. In this way, the classification accuracy was 88.5%, and the area under the receiver operating characteristic (ROC) was 0.8968.

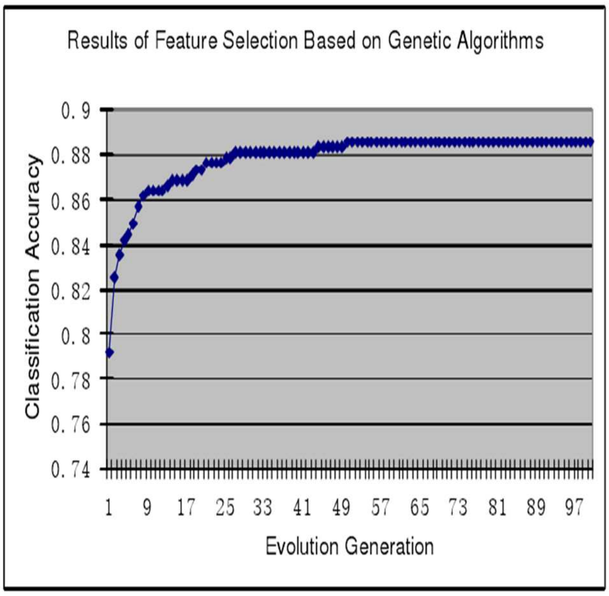

Figure 4. Results of feature selection based on genetic algorithm

T2I CNN classification

In addition to the traditional classification above, the T2I CNN classification was tested in this research. The new classification method aims to achieve computer feature exaction, and reduce the need for expert knowledge or manual intervention. To input the raw poems into the CNN structure, the cost function was defined and the network parameters were adjusted by stochastic gradient descent. For each input function, the cross-entropy function can be expressed as:

$$F=-(1/m)\Sigma y_i*(\log(h(x_i)))+(1-y_i)\log(1-h(x_i)) \quad (3)$$

where m is the total number of samples; x_i and y_i are the i-th labels of input x and the current sample y, respectively; h() can be derived from $x_0=w*x_i+b(1)$. Then, find the differentials of w and b in the structure by the cost function. The two parameters can be obtained by:

$$w= w+ 8(\partial f/\partial w) \quad (4)$$

$$b=b+(\partial f/\partial b) \quad (5)$$

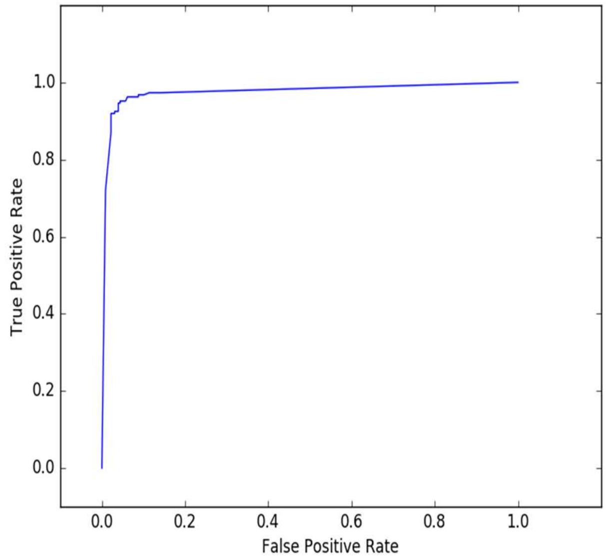

Figure 5. Results of the proposed algorithm

The CNN poetry style classifier was trained for 700 epochs. It can extract style features effectively without human interactions. Then, an experiment was performed on 5,600 labelled poems. The resultant ROC-enclosed area shown in figure 5 was 0.965, much greater than that of the genetic algorithm.

Results

For poetry style recognition, the styles of Ci-poetry were divided into Haofang and Wanyue. Several classification algorithms were employed to evaluate the style recognition results, including the naive Bayes and genetic algorithm with information gain, which is generated from expert selection and computer frequency counting. Then, the traditional methods were contrasted with the cognitive engineering method of T2I CNN classification. The results show that our method outperformed the traditional ones in both accuracy and ROC-enclosed area (Table 2). The advantages of our method are summarized as follows: First, less human expertise is required as the CNN imitates the human brain in parameter computing and makes connections just like real neurons; Second, the artificial neural network can handle big data raw materials, because it does not need the feature word list of genetic algorithm; Third, the T2I CNN poetry classification is more simple, direct and swift than genetic algorithms (e.g. it converges to a reliable and stable state in a few hours).

Table 2. Experiment Results and Evaluation

Method No.	Classification Algorithms	Accuracy	ROC
1	Naïve Bayes	0.502	0.394
2	Genetic Algorithm	0.885	0.8968
3	T2I CNN	0.912	0.965

Conclusions

The previous research on deep learning systems mainly emphasize on replicating the decision-making of human experts. However, the expertise is often inconsistent, unavailable and expensive. This is particularly true for studies in literature, like Classical Chinese Poetry. Different experts may hold exactly opposite views towards the same piece of poem. What is worse, the poetry style has become a matter of opinion. This calls for a uniform feature representation model that makes up for the defects of human expertise with computers.

To answer the call, the authors created a T2I CNN classification mode to migrate human brain connections to artificial neural networks. The central idea is to convert texts to image pixels, and input poems into the CNN. One of the main contributions is the successful implementation of a bi-directional character order input of the poem matrix, a common practice in bi-directional RNN structures for the NLP. Furthermore, the proposed method reduced the human labour on feature extraction of information gain in traditional genetic algorithms for poetry style recognition.

Nevertheless, both network layers and parameters are constrained by the available memory of GPUs in real experiments. Better classification can be achieved with faster GPUs, larger labelled poetry corpus, and more tolerance of training time. Based on poetry style recognition, the authors will incorporate brain poetry generation procedure into artificial neural network, creating smarter computer cognition of Classical Chinese Poetry.

References

Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence 2013; 35(8):1798-828.

Bengio Y, Ducharme R, Vincent P, Jauvin C. A neural probabilistic language model. Journal of machine Learning Research 2003; 3(Feb):1137-55.

Bird S, Klein E, Loper E. Natural Language Processing with Python; (J. Steele, Ed.) (1st ed.). Sebastopol, CA: O'Reilly Media, 2009.

Cadieu CF, Hong H, Yamins DL, Pinto N, Ardila D, Solomon EA, Majaj NJ, DiCarlo JJ. Deep neural networks rival the representation of primate IT cortex for core visual object recognition. PLoS Computational Biology 2014; 10(12):e1003963.

Cai H, Wang L, Duan S. Sentiment classification model based on word embedding and CNN. Application Research of Computers 2016; 33(10): 7-11.

Classical Chinese Poetry. (n.d.). In Wikipedia. Retrieved from https://en.wikipedia.org/wiki/Classical_Chinese_poetry (Retrieved on 26 August, 2017)

Collobert R, Weston J, Bottou L, Karlen M, Kavukcuoglu K, Kuksa P. Natural language processing (almost) from scratch. Journal of Machine Learning Research 2011;12:2493-537.

Connectionism 2015; Retrieved from https://plato.stanford.edu/entries/connectionism/

Corneli J, Jordanous A, Shepperd R, Llano MT, Misztal J, Colton S, Guckelsberger C. Computational Poetry Workshop: Making Sense of Work in Progress. Proceedings of the 6th International Conference on Computational Creativity 2015; 268–75.

Delmonte R. Computing Poetry Style. In C. Battaglino, C. Bosco, E. Cambria, R. Damiano, V. Patti, & P. Rosso (Eds.), Proceeding ESSEM - Emotion and Sentiment in Social and Expressive Media: approaches and perspectives from AI (ESSEM). Torino: CEUR Workshop Proceedings, 2013: 148-55.

Delmonte R. Exploring Shakespeare's Sonnets with SPARSAR. Linguistics and Literature Studies 2016; 4(1): 61–95.

Ge W, Yu Y. Borrowing treasures from the wealthy: Deep transfer learning through selective joint fine-tuning. InProc. IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, 2017; 6.

Hinton GE, Paccanaro A. Learning Distributed Representations of Concepts using Linear Relational Embedding. IEEE Transactions on Knowledge and Data Engineering 2001; 13(2): 232–44.

Hinton GE. Deep belief networks. Scholarpedia 2009; 4(5): 5947.

Hu R, Zhu Y. Automatic Classification of Tang Poetry Themes. Journal of Peking University (Science and Technology) 2015; 51(2): 262–68.

Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. In Advances In Neural Information Processing Systems 2012; 60(2): 1097–105.

Lecun Y, Bengio Y, Hinton GE. Deep Learning. Nature 2015; 521(7553): 436–44.

Li L, He Z, Yi Y. Poetry stylistic analysis technique based on term connections. Journal of Chinese Information Processing 2005; 19(6): 99–104.

Li L, He Z. General Model with Parameters Analysis of Syntax Tagging. Computer Science 2007; 34(11):189–92.

Li L. A study on term connection oriented NLP technique and its applications. Chongqing University, 2004.

Li L. Information dependency syntax tagging model. Shanghai: Academia Press, 2009.

Lin H. Toward Automated generation of chinese classic poetry. New Mathematics and Natural Computation. 2013; 9(02):153-81.

Liu C, Luo K. Tracking Words in Chinese Poetry of Tang and Song Dynasties with the China Biographical Database. arXive: Computation and Language 2016.

Liu CL. Quantitative analyses of Chinese poetry of Tang and Song dynasties: Using changing colors and innovative terms as examples. In Proc. of the International Confence on Digital Humanities 2013; 260–62.

Manurung R, Ritchie G, Thompson H. Using genetic algorithms to create meaningful poetic text. Journal of Experimental & Theoretical Artificial Intelligence 2012; 24(1):43-64.

Neurolinguistics. (n.d.). In Wikipedia. https://en.wikipedia.org/wiki/Neurolinguistics (Retrieved from 16 August, 2017)

Parkhi OM, Vedaldi A, Zisserman A. Deep Face Recognition. Proceedings of the British Machine Vision Conference 2015, (Section 3), 41.1–41.12.

Qian P, Huang X. Statistical Modeling and Macro Analysis on Chinese Classical Poems. Journal of Jiangxi Normal University (Natural Science) 2015; 39(2): 117–23.

Rahgozar A, Inkpen D. Poetry Chronological Classification: Hafez. In Canadian Conference on Artificial Intelligence. Springer, Cham 2016; 9673: 131–36.

Selangor B. Poetry Classification Using Support Vector Machines. Journal of Computer Science 2012; 8(9): 1441–46.

Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Sifre L. Mastering the game of Go without human knowledge. Nature 2017; 550: 354–59.

Srivastava N, Salakhutdinov R. Multimodal Learning with Deep Boltzmann Machines. In Advances in neural information processing systems (NIPS) 2012;2222–30.

Sundararajan L. Twenty-Four Poetic Moods: Poetry and Personality in Chinese Aesthetics. Creativity Research Journal 2004; 16(2-3): 201–14.

Wang Q, Luo T, Wang D. Can machine generate traditional Chinese poetry? A feigenbaum test. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics 2016; 10023 LNAI(100084): 34–46.

Wang Q. The Expressive Forms of Natural Imagery in Chinese Poetry. Advances in Literary Study 2017; 5(1): 17–21.

Weston J, Bengio S, Usunier N. Large scale image annotation: Learning to rank with joint word-image embeddings. Machine Learning 2010; 81(1): 21–35.

Yi Y, Zheng Y, He Z, Li L. Studies of Traditional Chinese Poet Identification Based on Machine Learning. Mind and Computation 2017; 1(60173060): 359–64.

Yi Y. A Study on Style Identification and Chinese Couplet Responses Oriented Computer Aided Poetry Composing; Chongqing University, 2005.

Yi X., Li R., Sun M. (2017) Generating Chinese Classical Poems with RNN Encoder-Decoder. In: Sun M., Wang X., Chang B., Xiong D. (eds) Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. CCL 2017, NLP-NABD 2017. Lecture Notes in Computer Science, vol 10565. Springer, Cham

Zhang X, Lapata M. Chinese Poetry Generation with Recurrent Neural Networks. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP'14) 2014; 670–80.

Zhou C, You W, Ding X. Genetic algorithm and its implementation of automatic generation of Chinese songci. Journal of Software 2010; 21(3): 427–37.