Research on the Model of Lyric Emotion Algorithm

Xiaoling Xia¹, Xin Gu²*, Qinyang Lu³
¹College of Computer Science, Donghua University
Shanghai, China
²College of Computer Science Donghua University
Shanghai, China
³College of Computer Science, Donghua University
Shanghai, China
*sherlysha@dhu.edu.cn b2171783@mail.dhu.edu.cn c2181818@mail.dhu.edu.cn

ABSTRACT: With the rapid development of Internet technology, various social networking platforms, especially mobile social networking platforms, continue to increase, resulting in a large amount of public opinion information. Internet public opinion has a clear emotional orientation, and its emotional orientation is very easy to spread and be infected, and even affect the development of the event. Aiming at the characteristics of lyric information rich and which are easy to change with time, the lyric theme analysis model and the lyric emotion evolution model are proposed. The LDA model is used to extract the topic from the lyric text in a period of time, and the sensational heat value is calculated according to the forwarding amount and the number of comments, and the lyrical theme with the highest heat is obtained. The relative entropy between sub-topics in the adjacent time slice of a specific hot topic is calculated, and the degree of association between the topics in the adjacent time slice is determined, thereby analyzing whether there is a split of the sub-topic and a new topic. Then the evaluation object is extracted, combined with the joint deep neural network model to judge the emotions of each evaluation object in different time, and the emotional evolution of the hot topic is analyzed from multiple dimensions. Finally, an example analysis of the network public opinion information from June to July 2018 is carried out to verify the validity of the above model. The model effectively solves the problems of immature emotion analysis model and low accuracy of emotion classification in the current public opinion analysis.

CCS Concepts
Computing methodologies → Neural networks

1. Introduction
Emotional analysis is to analyze people's opinions, emotions, attitudes toward certain products, services, and so on[1]. Emotional analysis has become one of the most active research areas in natural language processing. There has been a richer research on network public opinion. Hu et al.[2] used the topic model to study the latent semantics of public opinions. Studies have pointed out that, for a considerable extent, humans are very concerned about what other people think. For this reason, when humans need to make decisions, they often need to seek the opinions from others. Since the way to obtain information is mostly from the Internet and from social media, this leads to the public's emotions and emotions being easily affected by the one-sided information on the Internet. And when social media write news, it has a clear emotional orientation in its attitude, and its emotions are very
easy to spread and infect others, and even affect the development direction of the event. In recent years, on the Internet, there have been many re-shaping corporate images through social media, discussing star life, publishing opinions on news hotspots, and controlling people's emotions. There are more articles that affect social and political systems. The rapid development of social media and social networks has brought about a huge moral problem. Therefore, the current hot topic of social networks is extracted from the complicated and sensational information, and the emotional tendency of each hot topic is analyzed.[2] Extracting the trend of public opinion can further guide the correct direction of public opinion, stop the topic of socially unfavorable and maliciously guiding the public, and help the government monitor the emotional changes of the masses to avoid the occurrence of vicious incidents or false events. This is of great significance for ensuring social harmony and stability.

The most well-known way of natural language processing is based on both corpus[3] and dictionary-based[4][5] methods. The corpus is mainly based on the statistical characteristics of the large corpus, mining the evaluation words in the corpus and judging the polarity. The advantage lies in the simplicity and ease. The disadvantage is that the corpus has limited words, and the distribution of the evaluation words is not easy to generalize. Based on the dictionary mainly using the relationship between words and words to mine evaluation words, the difficulty lies in the need to update the dictionary to determine the meaning analysis. However, the lyric information is complex, and the traditional machine learning method[7] needs to extract features manually, ignoring the relevance of context between contexts, and it is difficult to obtain higher accuracy of sentiment classification. Boiy et al.[8] used a variety of machine learning methods such as maximum entropy model to mine emotional information. Ye et al.[9] compared different machine learning algorithms and found that the accuracy of using support vector machine and N-gram model is much higher than other learning algorithm; Wang et al[10]. segmented the data set according to the text topic, and separately trained the model of machine learning on the dataset of each topic, effectively improving the classification accuracy of the text. Therefore, how to extract the sentimental tendency from the Chinese lyric information is an urgent need to study and solve the problem in the field of lyric sentiment analysis. This paper uses the deep learning method to analyze the sentiment orientation of Chinese lyric information. The main research work is as follows:

Aiming at the characteristics of rich lyric information and rapid change with time, this paper proposes a lyric theme analysis model and a lyric emotion evolution model. According to the LDA model, the topic of the public opinion information acquired over a period of time is extracted. According to the social media correlation heat analysis, a plurality of lyrical topics with the best attention are obtained, and the relative degree of correlation between different topics is calculated by relative entropy, thereby obtaining the evolution of the topic. In the case, combined with the deep neural network model[11], the emotions of multiple topics are analyzed, and the emotional correlation and degree of association between different topics are obtained. Finally, experiments are carried out through a number of real lyric topics, which verifies the validity of the above model and effectively solves the problems of low classification accuracy of sentiment analysis and lack of public opinion evolution analysis model in current public opinion analysis.

2. RELARED WORK

Lyric emotion analysis model: The network has a rich theme, diverse content and fast update frequency. Therefore, the lack of time-sensitive lyric data is meaningless. The introduction of time stamps for lyric data can describe the lyric theme and the changes in emotions on the timeline, and can quickly find the center point of lyrics in different time periods. The public's sentiment towards each period. By comparing the emotional tendency and emotional development direction in different time periods, it is found that the relationship between public opinion highlights the importance of sentiment analysis, which is also the significance of public opinion analysis. This paper explores how to find hot topics from massive lyric information, applies sentiment analysis models to sentiment classification of different topics, introduces time concepts, and proposes lyric theme evolution models and lyric emotion evolution models to compare the trends of lyric topics and emotions on timelines.
**Lyric heat value calculation:** By crawling the lyric information of a certain amount of social networks, you can get the comment data, comment volume and praise number of each topic. The larger the number, the more the audience, the higher the attention and the wider the lyrics, so the sensation. Therefore, this article uses the number of comments and the number of likes to measure the heat value of the lyric topic.

According to the concept of information entropy, we generally use the logarithm of the inverse of the probability to represent the amount of information for an event. The more the information is determined, the larger the probability value $p$ is, and the smaller the value $I$ of the information entropy is. The formula is as follows:

$$I(p) = -\log p$$  \hspace{1cm} (1)

According to this formula, the formula for calculating the heat value of the public opinion can be derived. If the number of comments on the lyric topic is $c$ and the number of points is $L$, the heat value $h_m$ of the lyric $m$ is:

$$h_m = -\log\frac{1}{c+L+1}$$  \hspace{1cm} (2)

Where $c+L+1$ represents the sum of the number of comments and the number of likes and the number of authors. The higher the number of comments and the number of praises, the higher the degree of attention of the lyrics, the higher the heat value of the lyric topic. In order to obtain the topic popularity value in a period of time, it is necessary to model the topic of multiple lyric topics in the current time slice to obtain the topic distribution of each article. If the number of topics is $k$, then according to $k$ topics $\{z_1, z_2, ..., z_k\}$ contribution degree to judge the topic heat, then the topic $k$ is the hot value of the theme $k$:

$$H_{zk} = \sum_{i=1}^{m} p_{zk} \times h_i$$  \hspace{1cm} (3)

Where $H_{zk}$ is the heat of the subject $k$, which is obtained by weighted summation of the subject probability in the article and the heat value of the article. Finally, multiple lyric topics with the highest heat value and corresponding topic-word distribution are extracted, and the heat values of the same topic in different time slices are compared to analyze the evolution of the theme.

**Relative entropy calculation:** The evolution of lyric topics in social media is mainly divided into periods of germination, development, climax and recession, and may lead to new topics with the merger or division of certain topics. The calculation of the public opinion value mentioned above can only visually see the development and decline of different topics. It does not have an intuitive representation of whether the hot topic itself has merged or split. Therefore, this article focuses on the topic in different time. The similarity is measured, and it is thus judged whether a new topic has emerged within the hot topic. Relative entropy can be used to describe the difference between two probability distributions. The obtained public opinion data is divided into time slices, and the topic model is extracted for the public opinion data in different time slices to obtain the document-topic distribution and the theme-word distribution, and the theme in each adjacent time slice is recycled. The KL distance is calculated to determine whether the topics in different time slices are related and whether there is an evolution relationship.

By using the topic as a category tag, the topic-word distribution as a dataset, and using the average distribution difference between topics to measure the degree of dispersion of the lyric theme. If $k_1$ and $k_2$ are sub-topics in two adjacent time slices, $w_{i1}$ and $w_{i2}$ are the $i$-th words between the two sub-
topics, then \( p(W_i|k_1) \) is the i-th words’ probability distributions of \( k_1 \). \( p(W_i|k_2) \) is the probability distribution of the i-th word of \( k_2 \), and the formula for calculating the KL distance is:

\[
D(P(W_i|k_1)\mid P(W_i|k_2)) = \sum_i P(W_i|k_1) \log \frac{P(W_i|k_1)}{P(W_i|k_2)}
\]  

(4)

it can be known from the formula that if the two probability distributions are identical, the KL distance between them is zero. Therefore, a threshold can be set to determine the topic similarity within two time slices. When the KL distance of the adjacent time slice is smaller than the set threshold, it means that the topics in the two time slices may be related, and the heat of the previous topic does not disappear, indicating that there may be splitting within the topic; If the KL distance is greater than a given threshold, it indicates that the previous hot topic has declined and a new hot topic has been generated. Next chapter will design the public opinion evolution model from the two sub-models of the lyric theme evolution model and the lyric emotion evolution model.

**Lyric theme evolution model:** There are many research results in the topic classification model. The parameter distribution of the time label of the topic classification model mainly obeys the discrete distribution, obeys the continuous distribution, and converges according to the algorithm such as Markov chain. Obeying the continuous distribution of time label parameter settings can make the conditional distribution of parameters have continuity, the number of topics in each time slice is unchanged, but it will lead to an increase in similarity between topics, and the parameters in the pre-sequence time slice will impact on the current parameters causes the current lyric subject distribution information to be masked by the pre-order information, which has an impact on the model accuracy. Therefore, this paper uses a discrete-time topic model to analyze the migration of lyric topics, to ensure the independence of lyric information in each time slice, and to more effectively judge the emergence of new topics. The theme evolution model flow built in this chapter is as follows:

1. Climb the contents of Sina, NetEase, Tencent and other major portals and daily Weibo webpages for a period of time, including the posting of lyrics and lyrics, and the corpus data of the crawling to remove dryness and participles. After the pretreatment operation, the corpus D is obtained;
2. Disperse the corpus D into each time window with y as the time granularity. If it is finally divided into n time windows, then \( D = \{D_1, D_2, ..., D_n\} \);
3. For the corpus \( D_t = \{d_1, d_2, ..., d_m\} \) of the time slice \( t \), the document heat \( h_j = -\log \frac{1}{c+i+1} \) of \( d_j \) is calculated according to the number of comments and the amount of praise, and the document with higher heat is used. The lyric topic discussed is subject to more attention and discussion, Repeat this operation until \( j = m \);
4. Repeat step 3 for corpus D in each time slice until \( i=n \);
5. The LDA topic model training is performed on the corpus \( D_t = \{d_1, d_2, ..., d_m\} \) of the time slice \( t \), and the text-topic probability distribution is obtained by calculating the conjugate distribution of the theme and the Dirichlet and polynomial of the lyrics;
6. Through the document-topic probability distribution of the j-th document \( d_j \) and the calculation of the document heat \( h_j \), the heat value \( H_z \) of each topic \( z \) in the time slice \( t \) is obtained, and this operation is repeated until \( j=m \), and the same topic heat in the same corpus \( D_j \) is sought. And get the top p topics with the highest heat value;
7. Repeat steps 5 and 6 until \( i=n \);
8. Compare and analyze the p hot topics and heat values in different time slices \( t \), and get the heat history evolution of hot topics.
9. Fine-grained the time slice according to the specific situation, train the LDA model of the text of the adjacent time slice to obtain the subject-word probability distribution, cyclically calculate the KL distance of the sub-topics in the adjacent time slice, and judge whether the subject content has an evolution relationship. Get the results of the lyric theme evolution.
Lyric emotional evolution model: Considering that the theme of lyric information is closely related to lyric emotions, most netizens have a clear emotional tendency toward a particular topic when making comments. Therefore, this paper also introduces a lyric emotion evolution model. This model will focus on the characteristics of lyric information. Take into account the theme features of the lyrics. Many literature studies have shown that the combination of text features and external features related to this article can improve the learning efficiency of the model. For example, Wang Wei et al. use the topic model to discover key features in emotional information, integrate into the emotional vector space, and use machine learning algorithms to improve the accuracy of text classification. Ghosh inputs the text word vector and the grammatical feature vector of the text into the model together to obtain higher classification accuracy than before the fusion grammatical feature. Huang Faliang proposed to implement the theme and emotion through the LDA model with the synchronous derivation. This model will incorporate the topic information to which the word belongs as an enhanced feature into the text for emotional classification.

This article re-classifies the topic by citing the comment data to get the current text comment topic and classify the emotion. This method can more accurately get the comment on which aspect of the lyric information and what kind of emotion the comment has produced. The LDA model is used to classify the comment texts of each time slice, obtain the text core theme according to the text-topic distribution, enhance the text features according to the word-theme distribution, and input the text words together with the subject of the words into the joint neural network model, which is combined with CNN[12] and RNN[13]. Add external features to improve the classification effect, and finally get the theme and emotional tendency of the text. Finally, combined with the hot topics obtained in the lyric theme evolution model, the emotion classification of the corresponding theme is extracted for evolution analysis.

3. Experimental results and analysis
Lyric emotion analysis model: The network has a rich theme, diverse content and fast update frequency. Therefore, the lack of time-sensitive lyric data is meaningless. The introduction of time stamps for lyric data can describe the lyric theme and the changes in emotions on the timeline, and can quickly find the center point of lyrics in different time periods. The public's sentiment towards each period. By comparing the emotional tendency and emotional development direction in different time periods, it is found that the relationship between public opinion highlights the importance of sentiment analysis, which is also the significance of public opinion analysis. This paper explores how to find hot topics from massive lyric information, applies sentiment analysis models to sentiment classification of different topics, introduces time concepts, and proposes lyric theme evolution models and lyric emotion evolution models to compare the trends of lyric topics and emotions on timelines.

Experimental preparation: Because this experiment needs to divide the time slice in units of “days” and analyze the long-term public opinion information, combined with the characteristics of news websites and social networking sites, daily crawling of large news website data and daily hot news on Weibo Search data and corresponding comments and likes are used as analytical data for the lyric theme analysis model. This experiment climbed Sina News, NetEase News, Tencent News and Weibo Daily Hot Search from June 20 to July 30, 2018 as experimental corpus. The length of the news document is 200-1000 words, and the length of the comment data is about 50 words. The obtained comment data contains a lot of noise, so the lyric data needs to be preprocessed. 1. Remove a lot of duplicate data. 2. Remove short text of less than 5 words. 3. Deleting data that is not related to the subject classification and sentiment classification for the number, emoticon, and website link information existing in the web page.

Then, the corpus data is segmented, and the lyrics dictionary containing 2080 sentiment words is used for word segmentation, and finally the data after segmentation is deactivated. Data cleaning examples are shown in Table 1.

Lyric theme evolution model: Before performing LDA topic analysis on the text, it is necessary to determine the hyperparameters α and β of the LDA model and the optimal subject number K. At
present, the indicators for evaluating the performance of the topic model are mainly the degree of confusion. The degree of confusion is inversely proportional to the performance of the model. The smaller the confusion, the higher the efficiency, and the better the topic classification. If \(N_d\) represents the number of words contained in the text, \(w_{d,i}\) represents the i-th word in the text, then the confusion can be expressed as:

\[
\text{perplexity}(D) = \exp \left( -\frac{\sum_d \sum_i N_d \ln p(w_{d,i})}{\sum_d N_d} \right)
\] (5)

Firstly, based on experience, the approximate range of the number of model topics is determined, and then the models of different the number of topics are classified, so that the model with low confusion has good classification effect, and the number of topics corresponding to the model with low confusion can be determined as current. The optimal number of topics for the topic. After comparing the performance of the models under different themes, when the number of topics increases, the confusion decreases. When the number of topics increases to 10, the model confusion remains basically stable. Therefore, it is determined that the number \(K\) of the model is 10.

| Processing | Data |
|------------|------|
| Raw data   | # Barcas Wonder Kid # With the end of the last match between England and Colombia, all the top eight of the Russian World Cup came into being.: https://weibo.com/tv/v/jv5rtOcAY?fid=1034:4257941021485214 |
| Remove noise | # Barcas Wonder Kid # With the end of the last match between England and Colombia, all the top eight of the Russian World Cup came into being.: |
| Split word | /#/ Barcas Wonder Kid/# /With the end of/ the last match/ between/ England/ and/ Colombia/, /all the top eight/ of the/ Russian World Cup/ came into being.: |
| Stop words | /Barcas Wonder Kid/With the end of/ the last match/ between/ England/ and/ Colombia/, /all the top eight/ of the/ Russian World Cup/ came into being.: |

**Hot topic extraction:** According to the text in each time slice, the heat is calculated and the corresponding heat is marked. Then the preprocessed text is used to calculate the subject of each text in the LDA model, and the heat value \(H_{x_k}\) of each topic is calculated according to the heat calculation formula in the lyric theme evolution model. The hottest topic of each time slice is calculated by the heat value calculation. According to the output result of the fourth part, the topic content with the most public discussion in different time slices is obtained, and the keywords of the hottest topic in each time slice are selected for different time slices. The evolution of the theme of the theme. The change of hot topic keywords is shown in Figure 1.
It can be seen from the figure that the hotspots in each time slice are not similar. There are hot-spot topics that have not been reduced in hotspots, and there are also major public opinion events that burst out in various time films. The public opinion has a long-lasting influence on public opinion. The degree is higher, the lyrics spread more widely, and the lyrical topics that appear and disappear in the picture indicate that the public's attention is not high, or that the influence of public opinion is minimized under the control of relevant departments or relevant social media.

The following table (Table 2) shows the calculation of the first five topics with the highest time slice heat value and the first seven words describing the topic and the heat value corresponding to the topic.

According to the theme keywords, the hot spots that occurred during this period were: the knife-cutting event, the Russian World Cup, the girl jumping off the building, the 97th anniversary of the founding of the Communist Party of China, and the 40th anniversary of the reform and opening up. According to the facts, the hot events found in this model are consistent with the actual hot events, which proves that the validity of the topic heat value is calculated by the public opinion topic extraction and the text heat value.

Lyric emotional evolution model: According to the lyric emotion evolution model described in the previous chapter, this experiment will analyze the sentiment data emotionally. The texts of each comment text are obtained and the text of the hot topic is extracted. The Word2Vec model is used to train the word and the subject of the text word. The combination vector is input into the joint deep neural network for model training and application.

We manually label 10,000 lyric data, input the text vector and the theme vector into the model pair, train the model, compare the experimental results with the accuracy of the input text vector feature classification, and find that the text vector is combined with the theme vector. The high classification accuracy rate is more suitable for emotional classification of the lyric information with obvious feature features, which verifies the effectiveness of the feature by increasing the word theme.

The figure below (Figure 2) shows the accuracy comparison of the classification results of different input vectors.

| Topic       | Subject keyword                                      | Heat value $H_{ts}$ |
|-------------|------------------------------------------------------|---------------------|
| Topic 1     | person suspect Shanghai sword death sword cutting   | 99002               |
| Topic 2     | World Cup football German team Korea race eliminate  | 49869               |
| Topic 3     | jump off building obscenity Gansu idiotocia detentio n | 20232               |
| Topic 4     | The Communist Party liberate anniversary reform China July doctrine | 15507               |
| Topic 5     | coercion Shanxi rape filing crime deleted acceptance | 8501                |
Analysis of the evolution of lyric emotions: According to the lyric theme evolution model, a hot public opinion event was obtained within one month, and the lyrical content of the “Thai Shipwreck Event” was analyzed by a fine-grained time slice. This experiment will continue to use the lyric emotion evolution model to emotionally affect the event. Evolutionary analysis, emotional classification of lyric comments texts of different time slices, extraction of evaluation objects, and obtaining emotional classification results for each evaluation object of the topic. The figure below (Figure 3) shows the comparison of positive and negative sentiment classification results of the theme in time slice w.

It can be seen from the figure that in the first three days after the incident, the different aspects of the topic are generally positive emotions. Although the lyric attention is reduced in the last three days, the negative emotions begin to occupy a dominant position and the proportion of positive emotions decreases. The lyric emotion evolution model is subdivided for each type of topic, and the emotional tendency of each keyword in each topic is obtained. The positive and negative emotions of the 12 keywords under the theme of "Thai Shipwreck" are calculated, and the analysis is based on the analysis of lyric content. As a result, the top eight most representative keywords were combined and classified into four groups: “Thailand Capital”, “salvage saving”, “Haofeng Zhang hero”, and “death missing”. The following picture (Figure 4) shows the keywords in different time slices. The ratio of positive and negative sentimental tendencies.

Figure 2 Comparison of training results between different epoch models

Figure 3 Emotional analysis results of each time slice
In this experiment, an empirical study was carried out on the emotional evolution model of lyric emotions. Firstly, the esoteric events were analyzed and analyzed, and the hot topics were extracted. Then, the hotspot evolution analysis of each hotspot was carried out. After subdividing the time slice, the lyric content evolution analysis is carried out for a specific event, and the split time point of the sub-topic is obtained. In the lyric emotion evolution model, the validity of the input of the enhanced feature is improved to improve the lyric emotion classification result, and the sentiment analysis is combined with the real event to obtain the emotional change of each topic time point and the emotional tendency change of each split sub-topic.

4. Conclusion

This paper puts forward the lyric theme analysis model and the lyric emotion evolution model for the characteristics of lyric information which are rich and easy to change with time. The LDA model is used to extract the topic of the public opinion information from the lyric text in a period of time, and the lyrical heat value is calculated according to the forwarding amount and the number of comments, and the lyrical theme with the highest popularity is obtained, and the sub-topics of the adjacent hot topic are further selected for the specific hot topic. The relative entropy between the two is calculated to determine the degree of association between the topics in the adjacent time slices, thereby analyzing whether there is a split point of the sub-topic split and a generation point of the new topic. In order to judge the emotions on different sides of the same topic, this paper extracts the evaluation objects in different time, inputs the theme as the enhanced feature and the text feature into the joint deep neural network model, and carries out the emotional evolution of the hot topic from multiple dimensions.

Based on the characteristics of lyric texts, based on the improvement of traditional neural network model, this paper proposes a lyric theme analysis model and a lyric emotion evolution model to extract hot topics from heat, topic transfer and emotion evolution. Through the sentiment analysis of multiple evaluation objects, the public's emotional tendency to different aspects of the same hotspot can be more clearly understood, and it is of great significance to analyze and control the public opinion.

REFERENCES

[1] Dalong Xue., & Jun Ma. (2014). Online Public Opinion Analyst Course.
[2] Hu, Y., Bai, L., & Zhang, W. (2012). OLDA-based method for online topic evolution in network public opinion analysis. Journal of National University of Defense Technology, 34(1), 150-154.
[3] Yanyan Zhao., Bing Qin., & Ting Liu. (2010). Text Emotional Analysis. Journal of software, 21(8), 1834-1848.
[4] Liu B, Huang T, Cheng J. The automatic text classification method based on statistics. Journal of Chinese Information Processing 16(6), 18-24 (2002).
[5] Qin X, Yuan C A, Peng Y. Based on the dictionary method and genetic algorithm for text feature extraction. Computer Engineering and Design 29(21), 5651-5654 (2008).

[6] Xianghua, F., Guo, L., Yanyan, G., & Zhiqiang, W. (2013). Multi-aspect sentiment analysis for Chinese online social reviews based on topic modeling and HowNet lexicon. Knowledge-Based Systems, 37, 186-195.

[7] Pang B, Lee L, Vaithyanathan S. Thumbs up? sentiment classification using machine learning techniques[C]. EMNLP. Volume 10. 2002:79-86.

[8] Boiy E, Moens M F. A machine learning approach to sentiment analysis in multilingual Web texts[J]. Information Retrieval, 2009, 12(5):526-558.

[9] Ye, Q, Zhang Z and Law R. Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. Expert Systems with Applications, 2009. 36(3): P. 6527-6535.

[10] Wang X, Wei F, Liu X, et al. Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach[C]. International Conference on Information and Knowledge Management. ACM, Scotland, 2011:1031-1040.

[11] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[12] Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In European conference on computer vision (pp. 818-833). Springer, Cham.

[13] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.