Tan Jay

2022-09-15

Contents

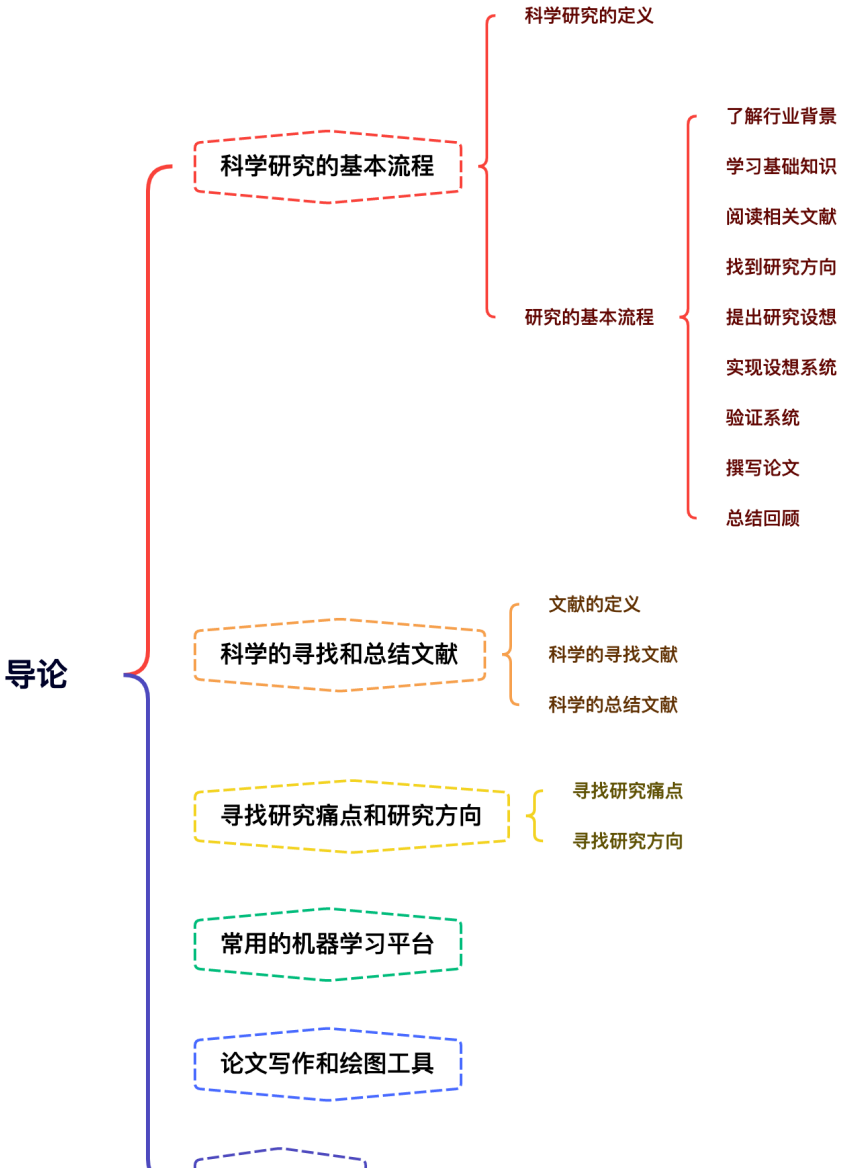
Chapter 1

bookdown

1. /
2. 6
3. 5
4. 3
5. 4
6. 7
7. 6
8.

Chapter 2

2.1


导论
科学研究的基本流程
科学研究的定义
研究的基本流程
了解行业背景
学习基础知识
阅读相关文献
找到研究方向
提出研究设想
实现设想系统
验证系统
撰写论文
总结回顾
科学的寻找和总结文献
文献的定义
科学的寻找文献
科学的总结文献
寻找研究痛点和研究方向
寻找研究痛点
寻找研究方向
常用的机器学习平台
论文写作和绘图工具

2.2

2.2.1

Research is “creative and systematic work undertaken to increase the stock of knowledge”. It involves the collection, organization, and analysis of information to increase understanding of a topic or issue. A research project may be an expansion on past work in the field. Research projects can be used to develop further knowledge on a topic, or for education.

2.2.2

1.

-
-

2.

-
-

3.

- https://scholar.google.com
- https://www.cnki.net
- IEEE Xplore http://ieeexplore.ieee.org
- ACM
- Elsevier
- Springer

4.

- NLP CV HCI
-
-

5.

-
-
-

6.

-
-
-
  -
  -
  -
  -
  -

7.

- / / 6
- 5
- 3
- 4
- 7
- 6
-

8.

-
-
-
-

2.3

-
-
-
-

2.3.1

1.

-
  -
  -
  -

  - GAN
  - SVM
-
  -
  -
  -
  -
-

2. +

- https://scholar.google.com
- https://www.cnki.net
- IEEE Xplore http://ieeexplore.ieee.org
- ACM
- Elsevier
- Springer

3.

-
-
- /

4.

-
  - CCF https://www.ccf.org.cn/
  -
-
  - : IEEE PAMI IEEE TNNLS Machine Learning
  - IEEE TIP IJCV
  - IEEE TASLP ACM SLP
  - IEEE Trans on Cybernetics IEEE TFS IEEE TAC IEEE TMI IEEE Trans on Robotics Pattern Recognition

2.3.2

1.

-
-
-

The three-dimensional (3-D) path planning of unmanned aerial vehicles (UAV) is a multi-objective optimization problem. It aims to find a smooth, flyable, and optimal path from starting point to the target point in a complex environment. Traditional algorithms have difficulty ensuring the optimal path

when faced with multiple objectives and complex cost functions. In this work, an improved butterfly optimization algorithm (IBOA) based on the virtual center butterfly (VCB) and Neighborhood dimension perturbation learning (NDPL) is proposed to solve the problem. Since BOA uses pairwise interactions between two random butterflies to perform the exploration. It makes the algorithm prone to miss the optimal solution, resulting in insufficient exploration capability. A novel VCB strategy was introduced into BOA to improve the exploration capability of the algorithm by creating attraction and repulsion effects on butterflies during the exploration phase. Meanwhile, the other individuals move toward the best individual in the exploitation phase, and if the best individual falls into the local extremes, it leads to premature convergence of the algorithm with low accuracy. A new NDPL is proposed, which constructs a neighborhood matrix after the BOA search is finished, then performs dimensional learning for each individual. The simulation experimental results in three scenarios show that the IBOA can acquire an effective and feasible route successfully, and its performance is superior to the other six algorithms.

-
- an improved butterfly optimization algorithm (IBOA)
- IBOA 6

2.

-
-
-
-

-
-
-
-

2.4

2.4.1

A.

B.

C.

D.

-
-
-
-

2.4.2

A.

-
-

1

1. CV ML
2.
3.

2

1. CV ML
2.
3.

B.

-
-
-

1

1.
2. CNN

2

1.
2.

C.

-
-

D.

-
-

2.5

- MATLAB
  - Matlab
  -
  -
- Caffe
  - C++
  -
  -
- PyTorch
  - Python
  -
  -
- TensorFlow
  - Python
  -
  -
- Karas
  - Python
  - API
  -
-
  - Theano
  - MXNET
  - Torch
  -

2.6

2.6.1

- Latex
- https://www.overleaf.com

2.6.2

- Microsoft Power Point
- Microsoft Power Visio
- OriginLab
- MATLAB
- Photoshop

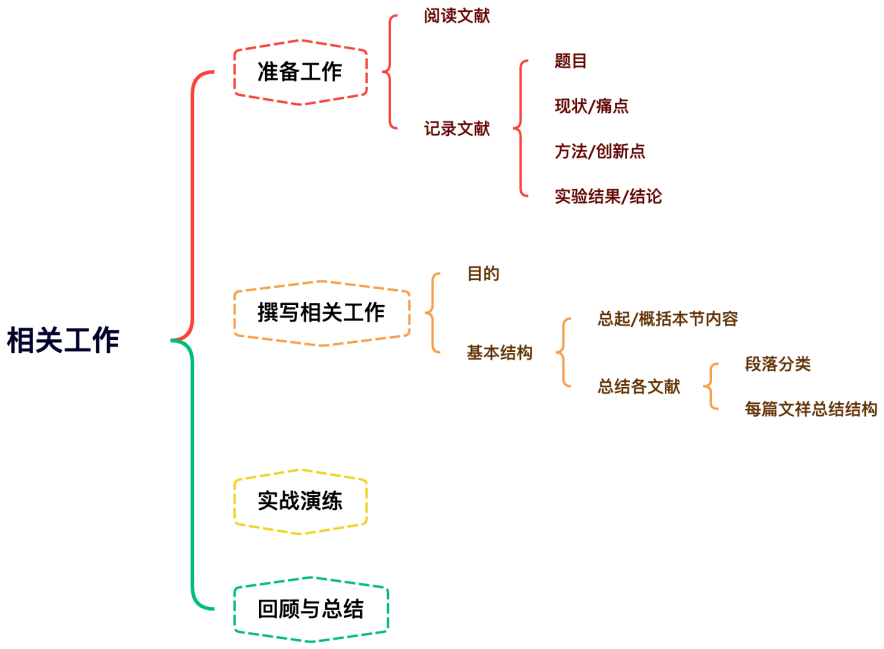
2.7

-
-
-
-
- Latex

Chapter 3

what why when how

3.1



3.2

Related Work means that the overall goal is to describe the related research areas and to place your method's contributions to the field in this context.

3.2.1

-
-
-
-

3.2.2

-
-
-
-

1

	Deep Residual Learning for Image Recognition The depth of representations is of central importance for many visual recognition tasks. Deeper neural networks are more difficult to train because of the vanishing/exploding gradient. While several solutions can deal with this issue, the accuracy of existing approaches still degrades rapidly with the networks depth increasing. Introducing a deep residual learning framework. The shortcut connections simply perform identity mapping. The results show that their approach allows deep network to be easier to train due to less complex strcuture and obtained the state-of-the-art performance on multiple image classification datasets

3.3

3.3.1

- The overall goal is to describe the **related research areas** and to place your method's contributions to the field in this context.
- By clearly describing previous work, you can **better describe the current limitations** and **the need for new methodology**.
- It also gives you an opportunity to demonstrate knowledge of the area and **helps others relate your current work to other scientific areas**.

3.3.2

- /
-
  –
  –

  /

1

- 1-2
  – In the past decade, many related studies have been proposed and published.
  – Since 1980, XXX ( ) has been rapidly developed.
-
  –

  These approaches can be mainly categorized into three classes:1.XXX;2.XXX;3.XXX.
  –

  Early studies mainly focused on XXX ( ), which XXX ( ). During XXX ( ), most researchers XXX ( ), but XXX ( ). In recent years, with the XXX ( ) becoming XXX ( ), this type of techniques have been widely employed in XXX ( ).

2

-
  – This section presents XXX in XXX. Then, XXX and XXX are introduced in XXX , which are followed by XXX.

  /

1. /
2. /
3. - / /
4. /

5.

- Albanie, Samuel, et al. “Emotion recognition in speech using cross-modal transfer in the wild.” Proceedings of the 26th ACM international conference on Multimedia. 2018.
  - Video as stack of still images
  - Video as spatial-temporal volumes
  - Short and Long-Term Dynamics
  - Multi-Stream Networks
  - Motion Information
  - Learning to Rank Videos
- Hu, Han, et al. “Relation networks for object detection.” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
  - Object relation in post processing
  - Sequential relation modeling
  - Human centered scenarios
  - Duplicate removal
- Junior, J. C. S.J., et al. “First impressions: A survey on computer vision-based apparent personality trait analysis.” IEEE Transactions on Affective Computing
  - The importance of first impressions in our lives
  - How challenging and subjective can be apparent personality trait labeling/evaluation?
  - Still Images
  - Image sequence
  - Audiovisual trait prediction
  - Multimodal trait prediction

--> --> / --> /

- Author proposed a method that XXX (/), which XXX (/).
- To deal with XXX ( ), XXX ( ) was employed by sb, which XXX ( ), and achieved XXX (/).

-

Biel et al. [17] studied personality impressions in conversational videos (vlogs) from facial expression analysis ( ). In their work, a subset of the Youtube vlog dataset [57] is adopted, as well as a facial expression model based on Facial Action Coding System (FACS) ( ). The task of first impressions prediction is

addressed using Support Vector Regression (SVR) combined with statistics of facial activity based on frame-by-frame estimates (1). Moreover, they analyzed what specific facial expressions are most prominent for modeling each of the different impressions (2) . Results show that extraversion is the trait showing the largest activity cue utilization, which is related to the evidence found in the literature that extraversion is typically easier to judge [58], [59] ( ). Later, Aran and Gatica-Perez [60] studied the use of social media content as a domain to learn apparent personality traits, in particular extraversion, aiming to transfer the knowledge extracted from conversational videos to small group settings ( ). Ridge Regression and SVM classifiers are combined with statistics extracted from weighted Motion Energy Images for the task of personality impressions prediction (/).

•

In low-level vision and computer graphics, for solving Partial Differential Equations (PDEs)( ), the widely used Multigrid method [3] reformulates the system as subproblems at multiple scales, where each subproblem is responsible for the residual solution between a coarser and a finer scale(/). An alternative to Multigrid is hierarchical basis pre-conditioning [44, 45], which relies on variables that represent residual vectors between two scales(/). It has been shown [3, 44, 45] that these solvers converge much faster than standard solvers that are unaware of the residual nature of the solutions ( ). These methods suggest that a good reformulation or preconditioning can simplify the optimization.

•

Practices and theories that lead to shortcut connections [2, 33, 48] have been studied for a long time( ). An early practice of training multi-laver perceptrons (MLPs) ( ) is to add a linear layer connected from the network input to thy, output [33, 48] ( ). In [43, 24], a few intermediate layers are directly connected to auxiliary classifiers for addressing vanishing/exploding gradients. ( ) The papers of [38, 37, 31, 46] propose methods for centering layer reponses, gradients, and propagated errors, implemented by shortcut connections. In [43], an inception" layer is com-posed of a shortcut branch and a few deeper branches. ( )

•
•
•

3.4

1.

	Deep Residual Learning for Image Recognition The depth of representations is of central importance for many visual recognition tasks. Deeper neural networks are more difficult to train because of the vanishing/exploding gradient. While several solutions can deal with this issue, the accuracy of existing approaches still degrades rapidly with the networks depth increasing. Introducing a deep residual learning framework. The shortcut connections simply perform identity mapping. The results show that their approach allows deep network to be easier to train due to less complex strcuture and obtained the state-of-the-art performance on multiple image classification datasets

2.

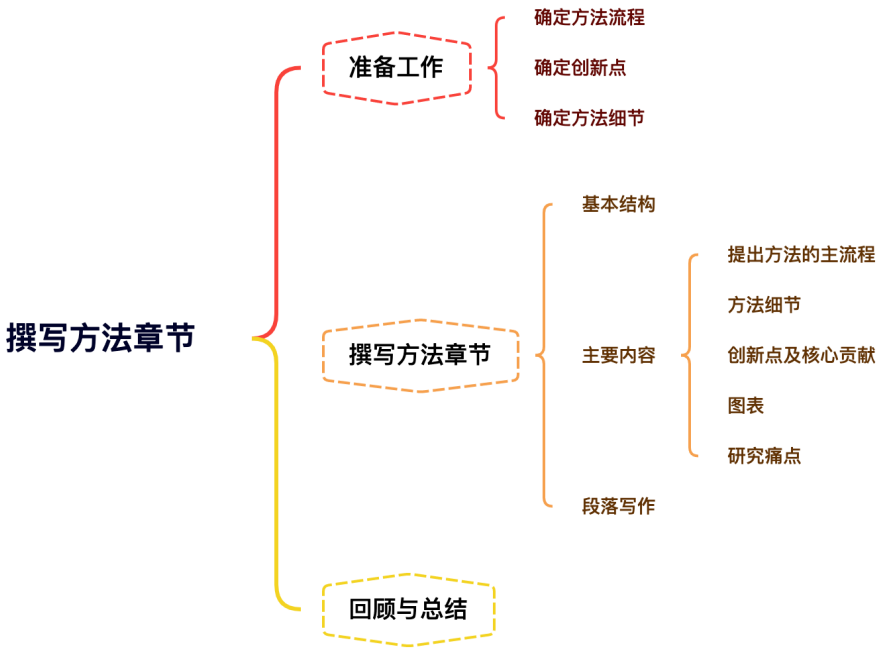
	To deal with the exploding gradient problem during the training as well as enhancing the performance of deep models. He et al. [1] proposed a deep residual network. The shortcut connections simply perfrom identity mapping, and their outputs are added to the outputs of the stacked layers. Obtained the state-of-the-art performance on multiple datasets and competitions.

3.

To deal with the exploding gradient problem during the training as well as enhancing the performance of deep models. He et al. [1] proposed a deep residual network, where the shortcut connections perform identity mapping, and their outputs are added to the outputs of the stacked layers. The results showed that this model obtained the state-of-the-art performance on multiple datasets and competitions.

Chapter 4

4.1



4.2

The proposed approach

The methods section of a research paper provides the information by which a study's validity is judged. The method section answers two main questions:

1. How was the data collected or generated?
2. How was it analyzed?

-
-
-

4.3

4.3.1

-
-
-

4.3.2

-
  -
  -
-
-
-
-

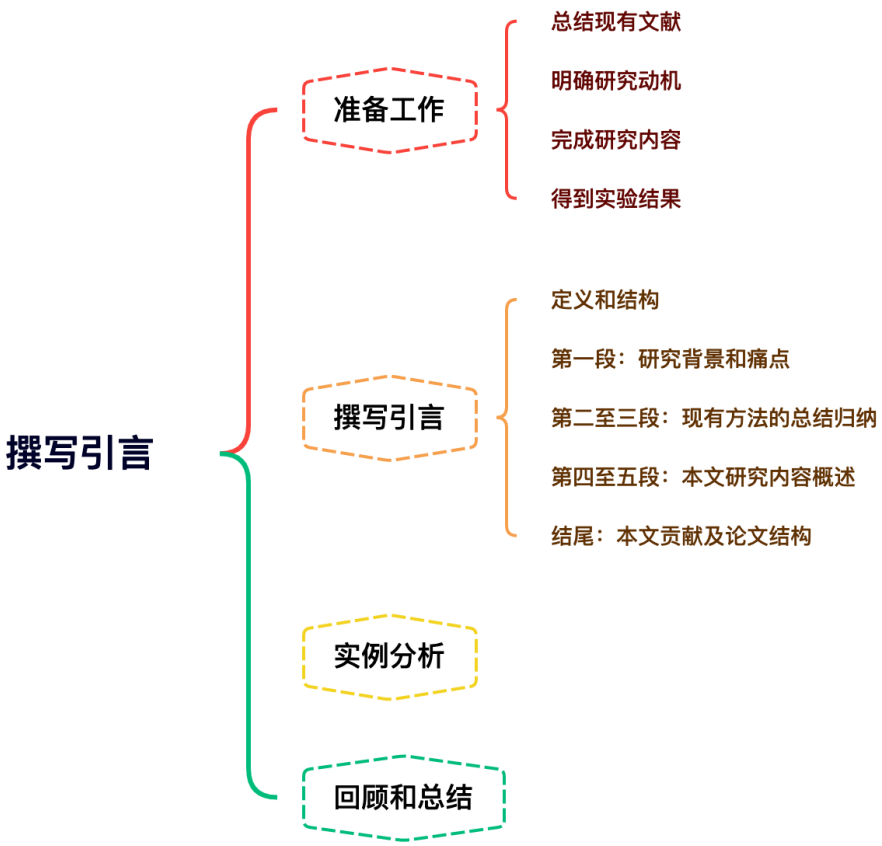
4.3.3

-
-

Chapter 5

what why when how

5.1



5.2

The Introduction section clarifies the motivation for the work presented and prepares readers for the structure of the paper.

1.

-
-
-

2.

-
-
-

3.

-
-
-

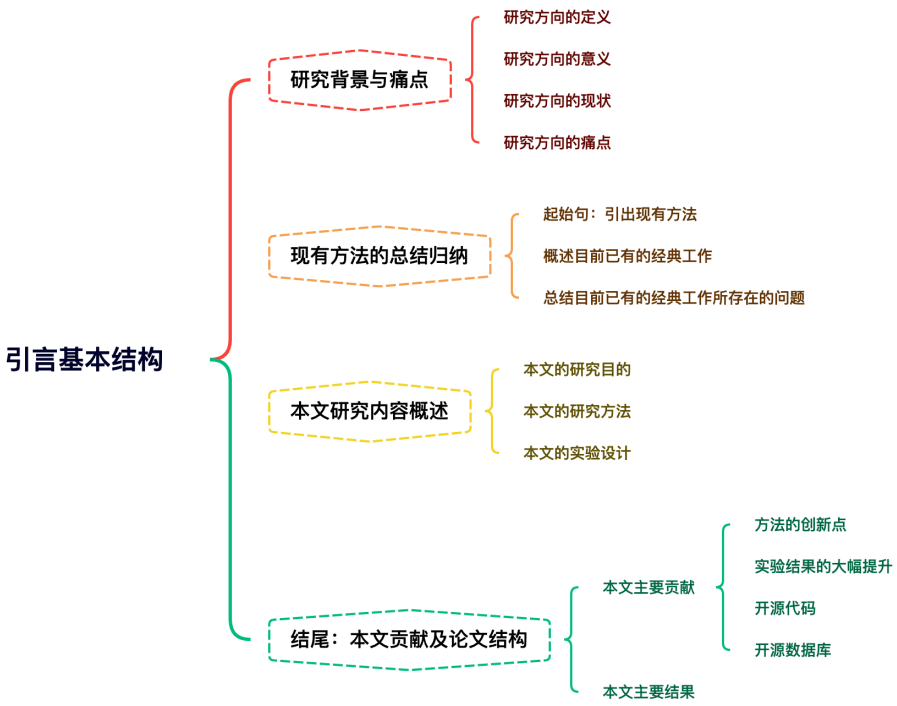
4.

-
-
- /

Introduction

5.3

5.3.1



5.3.2

-
-
-
-

5.3.3

1.

: - (With XXX becoming XXX/Due to the recent advance in), (a large number of attempts have been made/ a number of similar studies have been conducted) to XXX - XXX can be categorized into three fields: XXX, XXX and XXX - XXX is/are/becomes very popular in XXX field since XXX

2.

- :
- : (/ /)

- 1-2 (Related Work) ( )

In particular, most of these studies [6- 10] firstly extract hand-crafted visual descriptors from face images and then feed them to pre-trained classifiers for prediction ( ). Very early studies generally XXX (XXX) to obtain face representations, which XXX ( ). To solve these problems, Walker et al. [6] proposed a XXX method that (Walker). Alice et al. [7] XXX (Alice). However, both approaches XXX (1 Walker Alice). As a result, Luke et al. [8,9] and Yao et al. [10] specifically investigated XXX and ( ).

In particular, these studies [6-10] can be categorized into three categories: A-based approaches, B based approaches and C-based approaches ( ). The systems proposed by Luke et al. [8,9] and Yao et al. [10] are the typical examples of A-based approaches, which XXX (Luke Yao). Although A-based approach XXX (A), they XXX (A). Thus, B-based approaches have been widely investigated. For example, Alice et al. [7] XXX (Alice). Besides, Walker et al. [6] extends C to face recognition tasks, by XXX (Walker)

3.

Motivation. :

While the aforementioned approaches already achieved good recognition performance (around 90% accuracy) [9] on the face dataset that collected under controlled environments [11] they are still not able to provide reliable predictions on wild datasets [12-14] ( ). In addition, XXX ( ). This is because that existing hand-crafted visual descriptors are usually designed for general computer vision tasks without considering the face-specific information, which is crucial to face recognition tasks ( ).

5.3.4

1.

- To address/solve/deal with XXX, this paper presents/proposes XXX
- In this paper, we aims to XXX by XXX
- As a consequence, this paper XXX

2.

-
-
-

In this paper, we extend Convolution Neural Networks (CNNs) to the face domain (CNN). Specifically, the proposed approach starts with generating aligned face removing all background noises. Then, it feeds training faces to the proposed CNN model, paired with the corresponding labels. This way, the weights of the utilized CNN would be optimized, in order to correctly recognize the identity of the input face. In other words, the deep learned descriptors, which are generated by the well-trained CNN, are task -specific ( ). To the best of our knowledge, this is the first attempt to extract deep face descriptor for face recognition task ( ).

3.

-
-
-

To evaluate the performance of the proposed deep learning approach on both controlled and wild conditions, this paper conducts experiments on XXX dataset and XXX dataset ( ). In addition, we also evaluate the learned deep feature on XXX face tasks ( ). The experimental results show that ( ).

4.

-
-
-
-

5.3.5

1.

The main contributions of this study can be summarized as follows: This paper proposes the first deep learning-based face recognition system, which learns task-specific face descriptors rather than hand-crafted general visual descriptors ( ). The proposed approach achieved the state-of-the-art XXX performance on XXX dataset and XXX dataset, clearly outperform other existing approaches ( ). The code of the paper is made publicity available at XXX ( ).

2.

The rest of the paper is organized as follows. In Sec. 2, we describe . XX are then presented in Sect. 3, and the XXX are presented/reported in Sect. 4. Sect. 5 concludes the paper. For completeness, we also describe and assess XXX in Appendix A.

5.4

Face recognition system is a technology capable of identifying or verifying a person from a digital image or a video frame from a video source ( ). Face recognition are preliminary steps to a wide range of applications such as personal identity verification [1, 2], video-surveillance [3], lip tracking [4,5], facial expression extraction [6-8], etc ( ). While there is a large number of studies that have been devoted to the such field, existing approaches generally employed various hand-crafted visual descriptors and generated relatively low and unstable recognition performance on wild datasets ( ).

Due to the recent advance in hand-craft features and classifiers, a large number of attempts have been made to face recognition area. In particular, these studies [6-10] can be categorized into three categories: A-based approaches, B-based approaches and C-based approaches ( ). The systems proposed by Luke et al. [8,9] and Yao et al. [10] are the typical examples of A-based approaches, which XXX (Luke Yao). Although A-based approach XXX (A) they XXX (A). Thus, B- based approaches have been widely investigated. For example, Alice et al. [7] XXX (Alice). Besides, Walker et al. [6] extends C to face recognition tasks, by XXX (Walker).

While the aforementioned approaches already achieved good recognition performance (around 90% accuracy) [9] on the face dataset that collected under controlled environments [11] they are still not able to provide reliable predictions on wild datasets [12-14] ( ). In addition, XXX ( ). This is because that existing hand-crafted visual descriptors are usually designed for general computer vision tasks without considering the face-specific information, which is crucial to face recognition tasks ( ).

As a consequence, this paper proposes to extract task- specified descriptors for face recognition, aiming to further enhance the face recognition performance under both controlled and wild conditions. In this paper, we extend Convolution Neural Networks (CNNs) to the face domain (CNN). Specifically, the

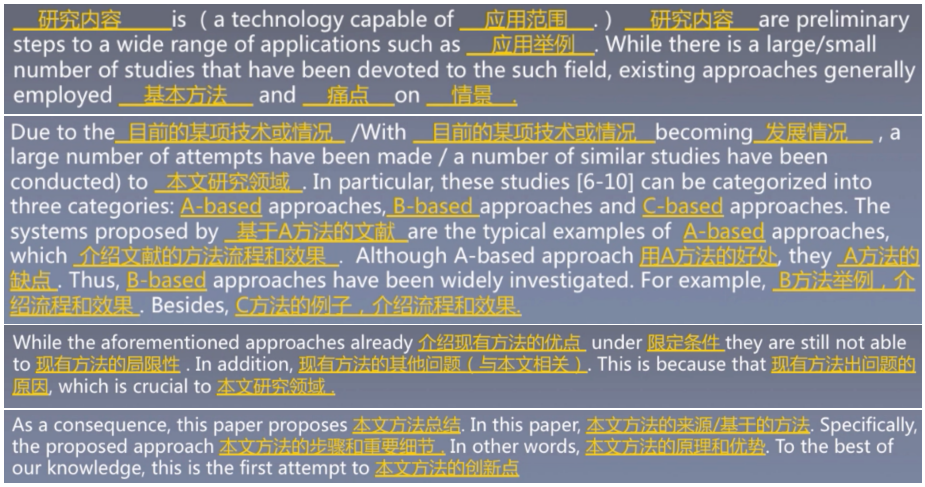
proposed approach starts with generating aligned face removing all background noises. Then, it feeds training faces to the proposed CNN model, paired with the corresponding labels. This way, the weights of the utilized CNN would be optimized, in order to correctly recognize the identity of the input face. In other words, the deep learned descriptors, which are generated by the well-trained CNN, are task-specific ( ). To the best of our knowledge, this is the first attempt to extract deep face descriptor for face recognition task ( ).

To evaluate the performance of the proposed deep learning approach on both controlled and wild conditions, this paper conducts experiments on XXX dataset and XXX dataset ( ). In addition, we also evaluate the learned deep feature on XXX face tasks ( ). The experimental results show that ( ). The main contributions of this study can be summarized as follows: The main contributions of this study can be summarized as follows:

1. This paper proposes the first deep learning-based face recognition system, which learns task-specific face descriptors rather than hand-crafted general visual descriptors ( ).
2. The proposed approach achieved the state-of-the-art XXX performance on XXX dataset and XXX dataset, clearly outperform other existing approaches ( ).
3. The code of the paper is made publicity available at XXX ( ).

The rest of the paper is organized as follows. In Sec. 2, we describe . XX are then presented in Sect. 3, and the XXX are presented/reported in Sect. 4. Sect. 5 concludes the paper. For completeness, we also describe and assess XXX in Appendix A.

5.5



To evaluate the performance of the proposed 提出的方法 approach on 应用场景, this paper conducts experiments on 实验数据库/实验场景 . In addition, we also evaluate the 提出的方法 on 另一个实验场景. The experimental results show that 简要分析实验结果. The main contributions of this study can be summarized as follows:

1. 方法上的创新
2. 好的实验结果
3. 开源代码或数据库

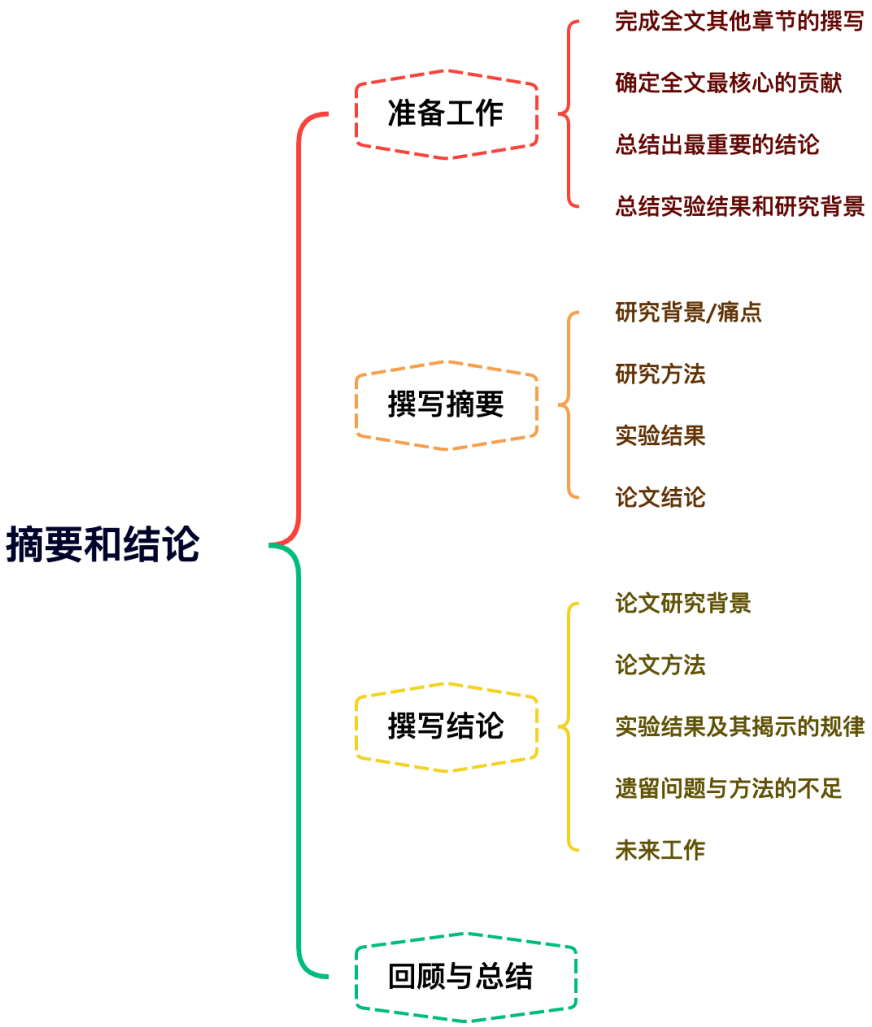
The rest of the paper is organized as follows. In Sec. 2, we systematically review previous face recognition studies. The proposed deep learning-based face recognition approach are then presented in Sec. 3, and the detailed experimental settings and results are reported in Sec. 4. Sect. 5 concludes the paper. For completeness, we also describe and assess the computation complexity in Appendix A.

5.6

-
-
-

Chapter 6

6.1


摘要和结论
准备工作
完成全文其他章节的撰写
确定全文最核心的贡献
总结出最重要的结论
总结实验结果和研究背景
撰写摘要
研究背景/痛点
研究方法
实验结果
论文结论
撰写结论
论文研究背景
论文方法
实验结果及其揭示的规律
遗留问题与方法的不足
未来工作
回顾与总结

6.2

A|

-
- / ( )
- /
-

6.3

- / (1-2)
- (2-4)
- (1-2)
- (1)

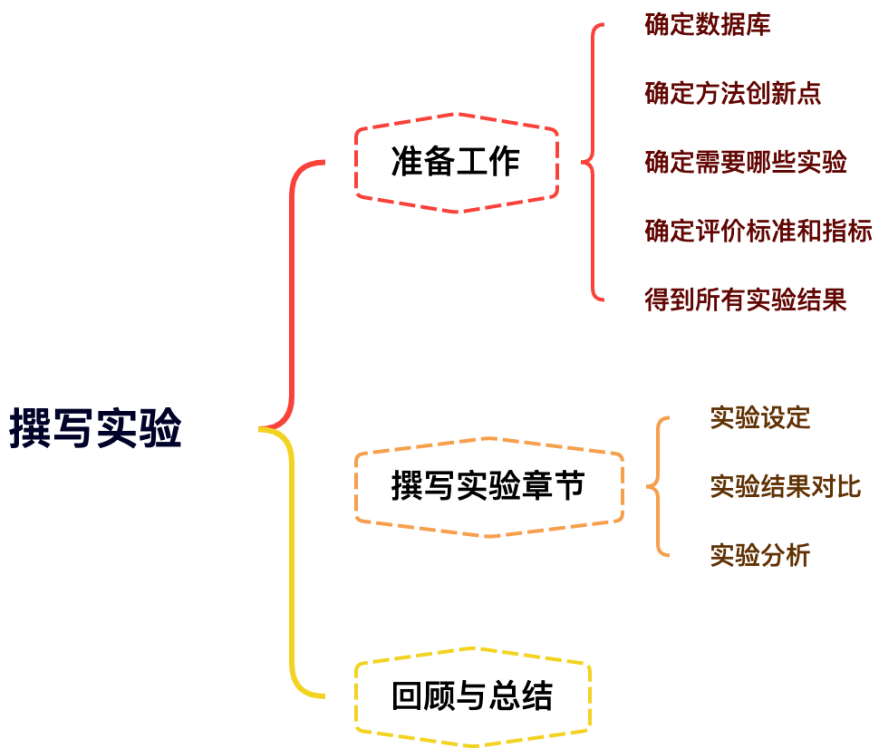
6.4

-
-
-
-
-
-

- / (1)
- (2)
- (1-2)
- (3-5)
- (2-3)

Chapter 7

7.1


撰写实验
准备工作
确定数据库
确定方法创新点
确定需要哪些实验
确定评价标准和指标
得到所有实验结果
撰写实验章节
实验设定
实验结果对比
实验分析
回顾与总结

7.2

This section should present and discuss the research results, respectively. However, because readers can seldom make sense of results alone without accompanying interpretation they need to be told what the results mean.

-
  -
  -
  -
  -
-
-
  -
  -
  - / /
-
  - Recall, Precision, F1, Accuracy
  - MAE, MSE, PCC, CCC
  - Inception Score
- :
  -
  - :
  - :

7.3

-
-
-

7.3.1

-
-
-
- tricks

7.3.2

- / (Ablation studies)
-

7.3.3

-
-
-

7.4

- (Database)
- (Implementation details)
- (Evaluation Metrics)
- (Experimental Results)
  - Ablation studies
  - Comparison to others
- (Result analysis/Discussions)