



基于迁移学习的广义 地图信息量计算

汇报人：王成龙



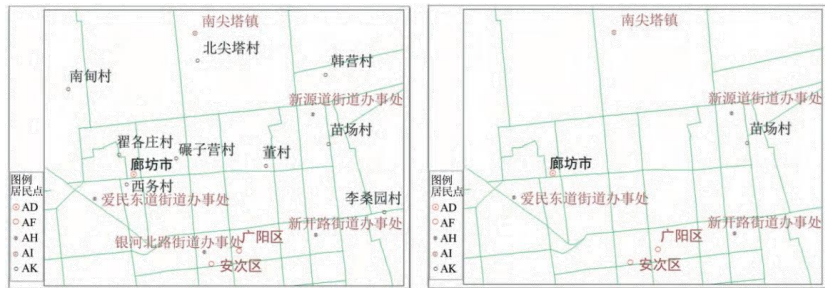
目 录

CONTENTS

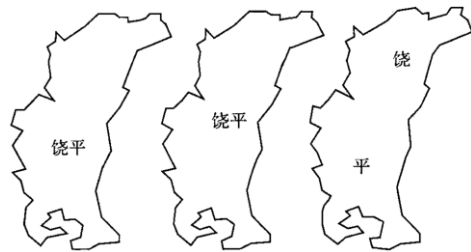
- 01 需求与研究概述
- 02 研究基础与现状
- 03 研究范围的界定
- 04 数据与模型构建
- 05 结果与未来展望

实际需求

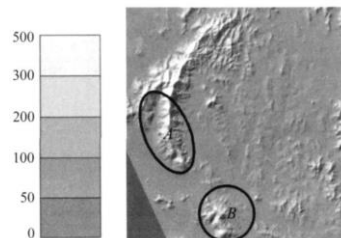
- 地图设计/制图综合/地图评价/...



(a) 要素选取[数量与等级]



(b) 注记配置[空间结构]



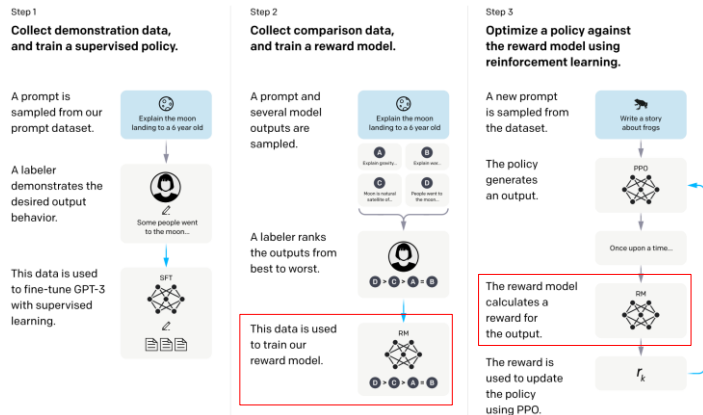
(c) 分层设色[色彩]

研究概述

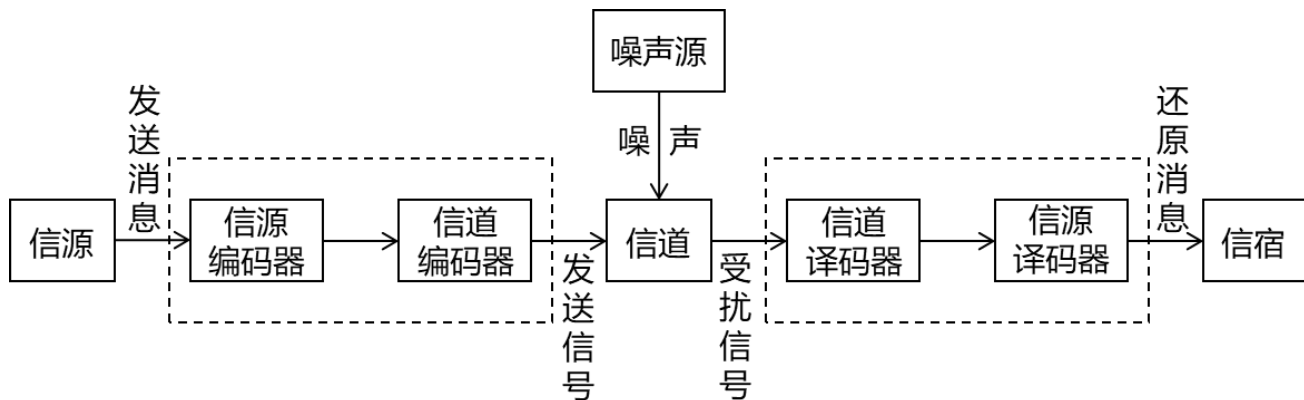
- 目标：自动度量广义信息(狭义信息+主观感受)



- 前景：地图自动编制(奖励模型)



InstructGPT模型架构



信息·论域

语用信息：事物运动状态及变化方式对观察者的效用

语义信息：事物运动状态及变化方式的含义

语法信息：事物运动状态及变化方式的形式
(剥离语义剩下的不确定性)

通讯·目的与核心问题

目的：语法信息的传递和交换

核心问题：信息传输的有效性和可靠性以及两者间的关系

信息论·度量

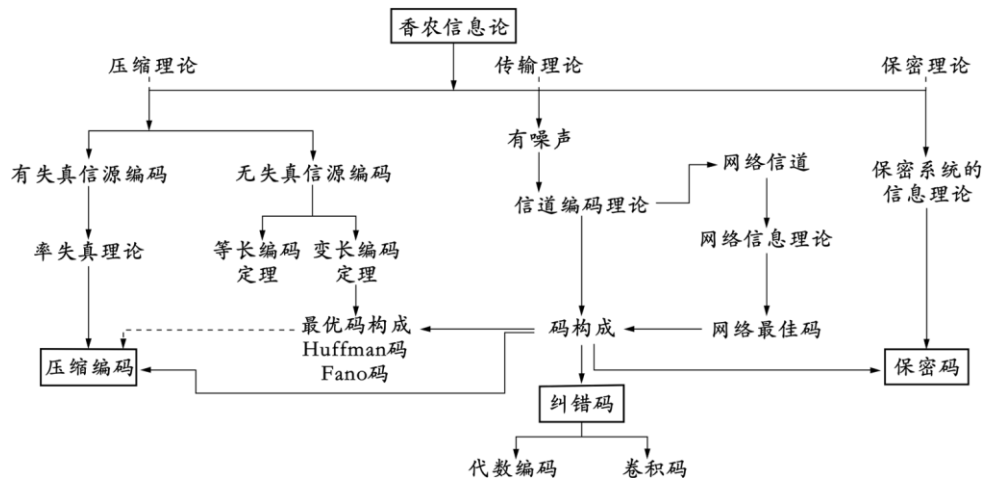
自信息：表示每个消息提供的平均信息量

$$I(x_i) = \log \frac{1}{P(x_i)}$$

信息熵：信源所包含的平均信息量，表征信源输出后每个信息提供的平均信息量

$$H = - \sum_i P(x_i) \log P(x_i)$$

式中, $P(x_i)$ 表示随机变量 X 每一个可能取值 x_i 的概率



信息论·三大定理

无失真信源编码定理

有噪信道编码定理

保失真度准则下的有失真信源编码定理

信息论·关注点

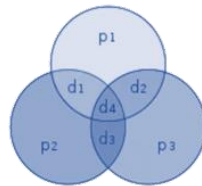
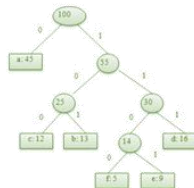
信源与信道

数据压缩、数据传输、密码学等

信息论·最为常见应用

信源编码：Hoffman编码

信道编码：圣经中的校验码/Hamming编码



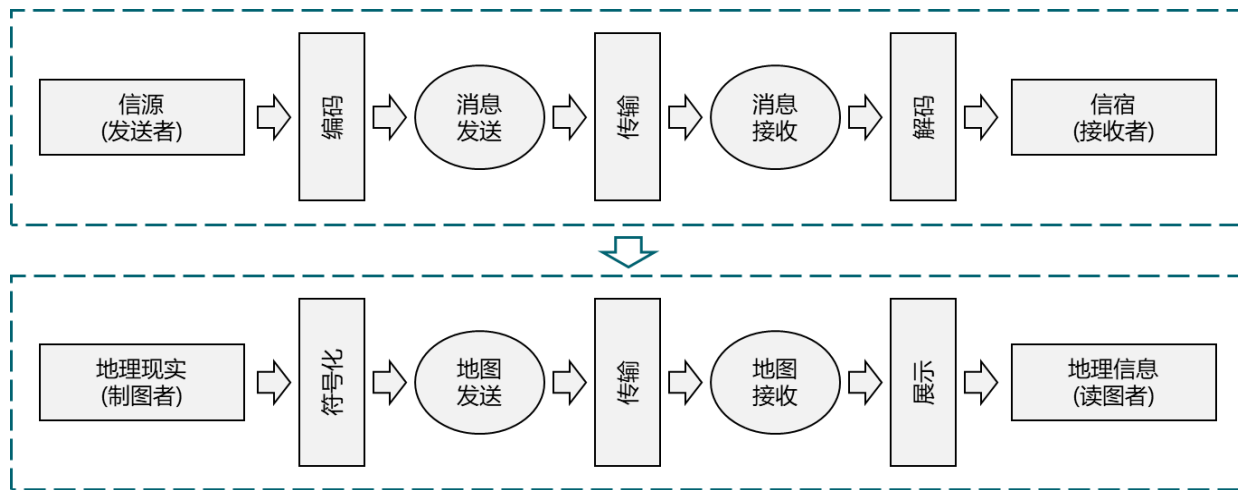


地图信息论·必要性

地图被认为是人类最古老的3大交流工具之一

地图的基本功能包括：模拟功能、信息负载功能、信息传输功能和认知功能等

地图信息论·启示

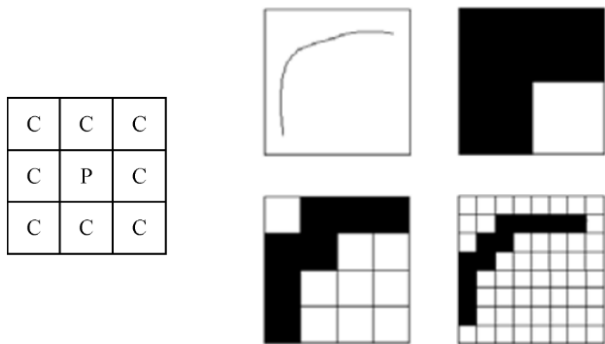




• 研究现状

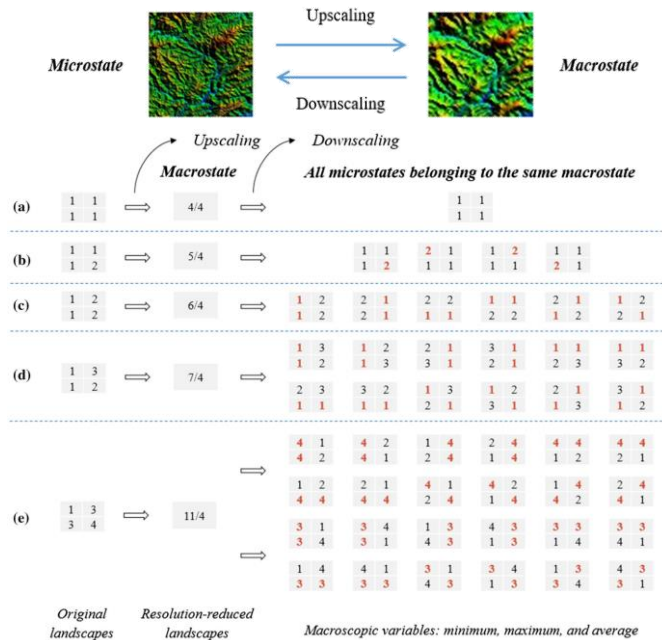
- 基于灰度差的方法(左)
- 基于玻尔兹曼熵的方法(右)
- 基于狭义/广义信息熵的方法

图像(Pixel-level)还是地图(Graphic-level)信息量?



$$Information_p = \sum (Difference\ between\ P\ and\ its\ context)$$

(a)基于灰度差的方法

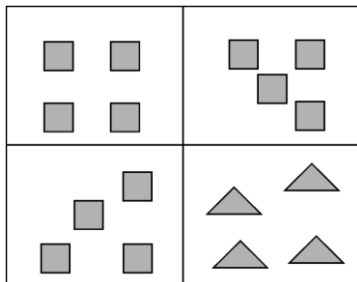


(b)基于玻尔兹曼熵的方法

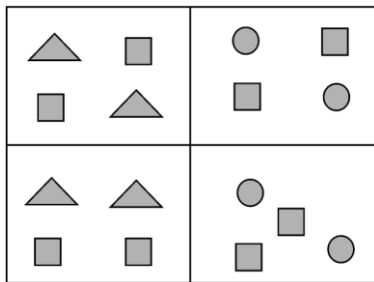


• 研究现状

- 基于灰度差的方法
- 基于玻尔兹曼熵的方法
- 基于狭义(左)/广义(右)信息熵的方法



(a) 4个符号一样, 信息等于零



(b) 2类符号(各2个)信息量为0.3

$$P(x_i) = \frac{K_i}{N} \quad i = 1, 2, \dots, M \quad -2 \times \frac{1}{2} \log_2 \frac{1}{2} = 0.3$$

(a) 基于狭义信息熵的方法

信息熵基于统计, 地图信息包含空间结构

度量信息

与位置/大小/分布/空间取向等有关

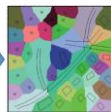
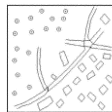
主题信息

与类别/重要性等有关

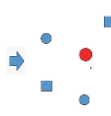
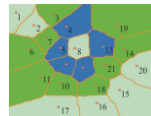
关系信息

与拓扑/距离/次序等有关

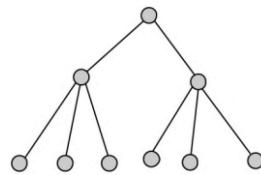
$$H(M) = - \sum_{i=1}^N \left(\frac{S_i}{S} \right) \times \log \left(\frac{S_i}{S} \right)$$



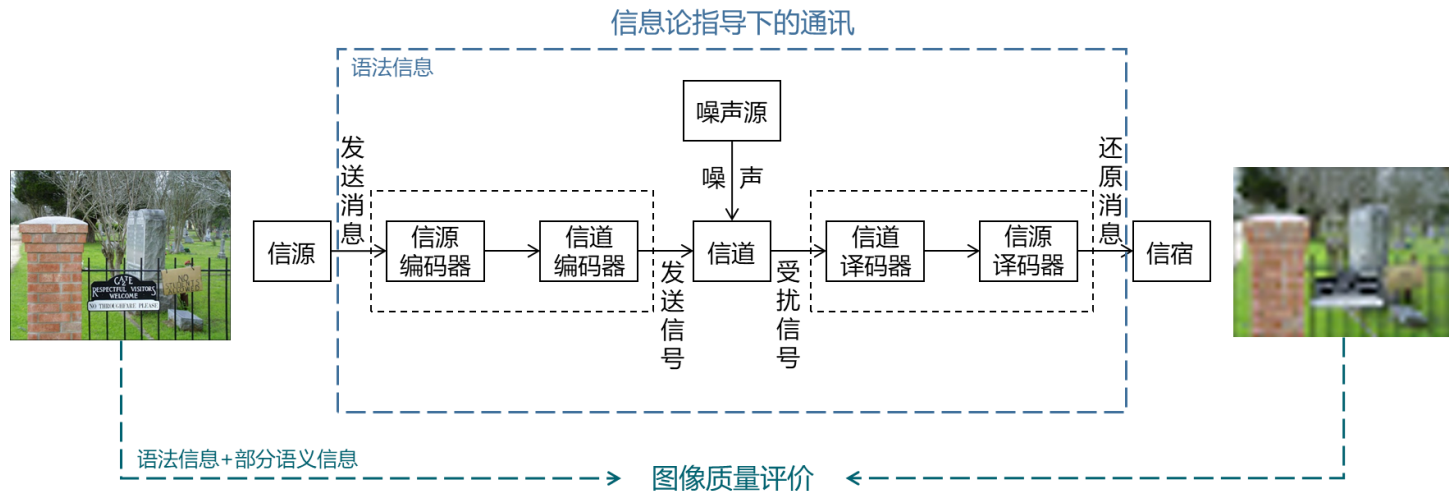
$$H(T) = \sum_{i=1}^N H_i(T) = - \sum_{i=1}^N \sum_{j=1}^{M_i} \left\{ \frac{n_{ij}}{N_i} \log \frac{n_{ij}}{N_i} \right\}$$



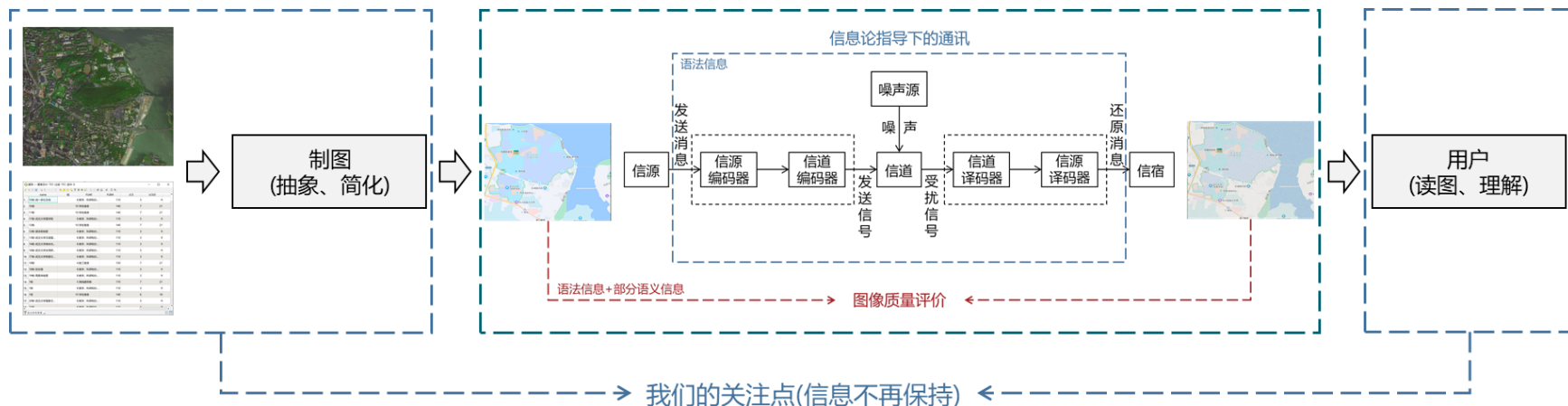
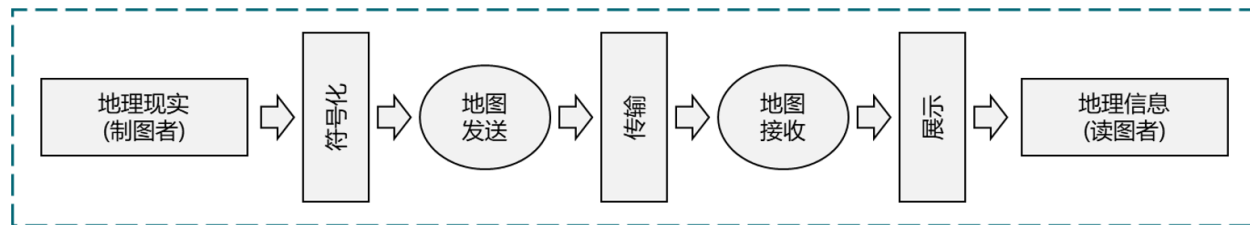
(b) 基于广义信息熵的方法



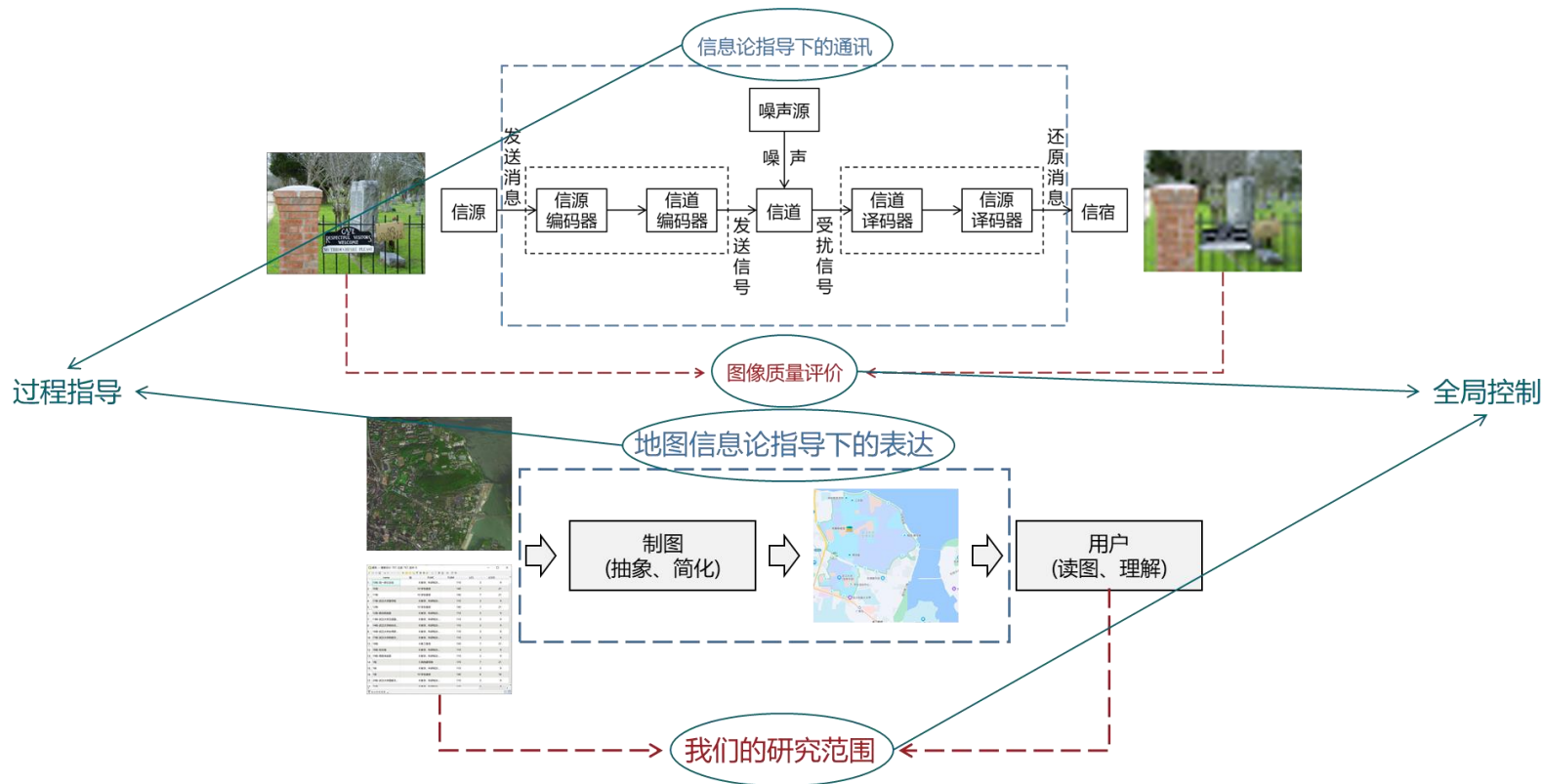
仅适用于矢量而不适用于栅格; 没有考虑人的主观感受
时间复杂度高; 关系信息度量结果非熵



图像质量评估(Image Quality Assessment, IQA)是一种用于自动预测图像感知质量的技术。IQA估计值应与一系列人类评估人员进行的质量评估高度相关(Mean Opinion Score, MOS)。



地图信息论建立在经典信息论基础之上，以图像为载体，关注点不再是传输，而是有效表达



常见指标: PSNR/SSIM/...(逐渐考虑人类视觉系统与感受)

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX^2}{MSE} \right)$$

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [I(i,j) - K(i,j)]^2$$

Scores: 36.9

Scores: 26.7

...

Scores: 60.0

Dataset	Reference Img. No.	Distorted Img. No.
TID2013	25	3000
TID2008	25	1700
CSIQ	30	866
LIVE	29	779
IVC	10	185
Toyama	14	168
A57	3	54
WIQ	7	80

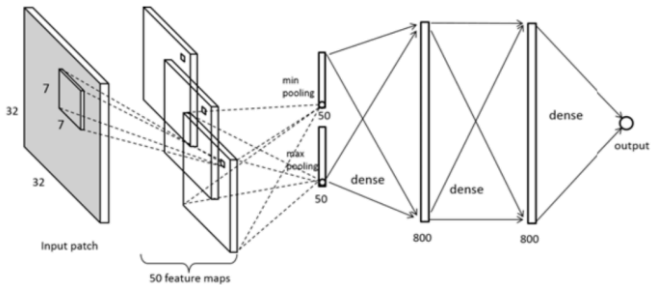
$$SROCC = 1 - \frac{6 \sum_{i=1}^N (v_i - p_i)^2}{N(N^2 - 1)} \quad LCC = \frac{\sum_{i=1}^N (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^N (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2}}$$

图像质量评价·无参图像质量评价(方法一:构造特征)

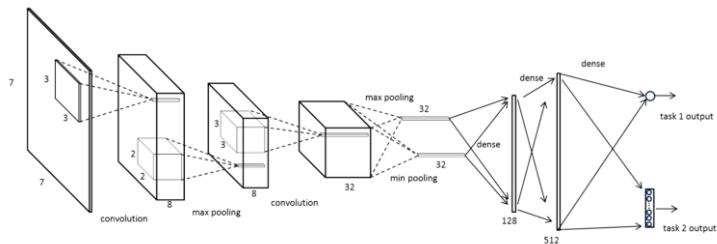
图像质量评价·无参图像质量评价(方法二:基于深度学习)

核心问题: 小样本训练

解决方案一: 切块扩大样本量



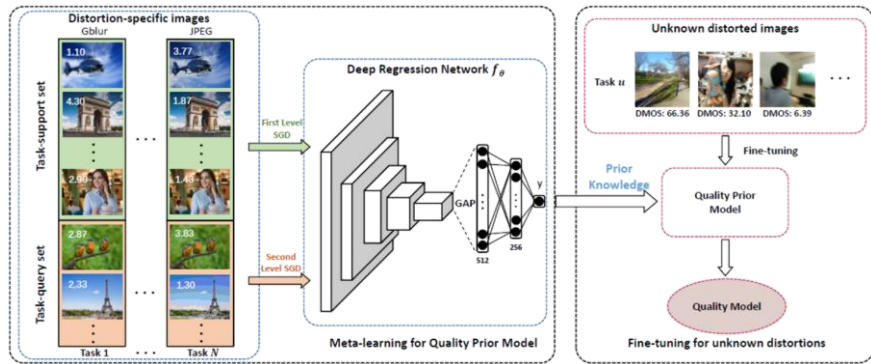
解决方案二: 多任务学习



解决方案三: 迁移学习

	LCC	SROCC	nMAE
Imagenet-CNN	0.6782	0.6381	0.12
Places-CNN	0.6267	0.6055	0.12
ImageNet+Places-CNN	0.7215	0.7021	0.11

解决方案四: 元学习



- **类比全参质量评价**
 - 将成果与原始数据进行比较来判断制图表达的好坏
- **类比无参质量评价**
 - 仅根据制图结果评价地图广义信息量与视觉冲击



图例	名称	面积	比例	备注
1	水域	150	5	
2	林地	150	5	
3	草地	150	5	
4	耕地	150	5	
5	建设用地	150	5	
6	未利用地	150	5	
7	其他	150	5	
8	总计	150	5	



Scores: 36.9

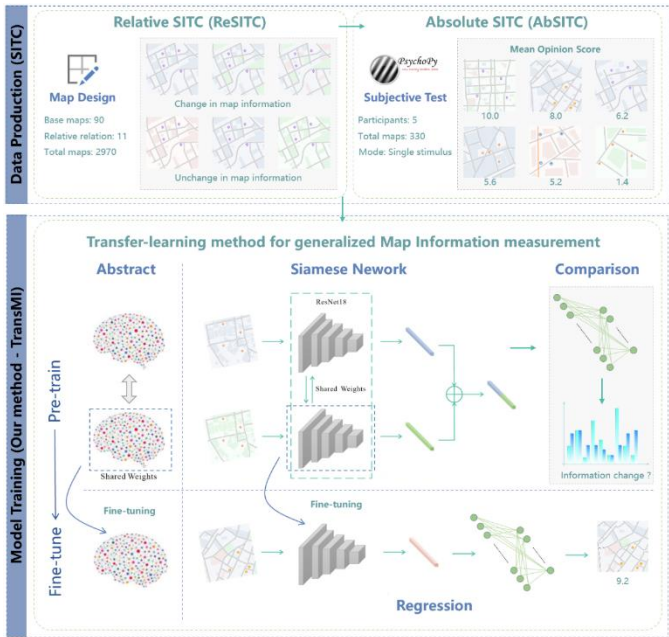


Scores: 60.0



Scores: 26.7

- **研究工作**
 - 构建地图信息论领域第一个适用深度学习的主观数据集
 - 基于孪生神经网络创新性提出一种全新的地图信息计算框架



绝对数值数据集

- 330张地图邀请5位地图与地理信息专业学者打分



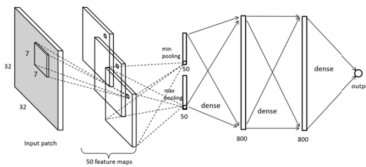
使用视觉模型训练

- 模型表现不佳(仍然面临小样本学习问题)

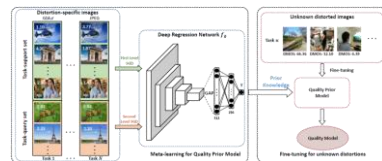
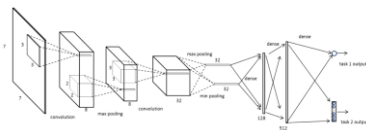


挑战

- 数据增广(信息量并不存在局部-全局不变性)
- 多任务学习(无其他代理任务)
- 元学习(无任务相关数据集)



	LCC	SROCC	nMAE
Imagenet-CNN	0.6782	0.6381	0.12
Places-CNN	0.6267	0.6055	0.12
ImageNet+Places-CNN	0.7215	0.7021	0.11



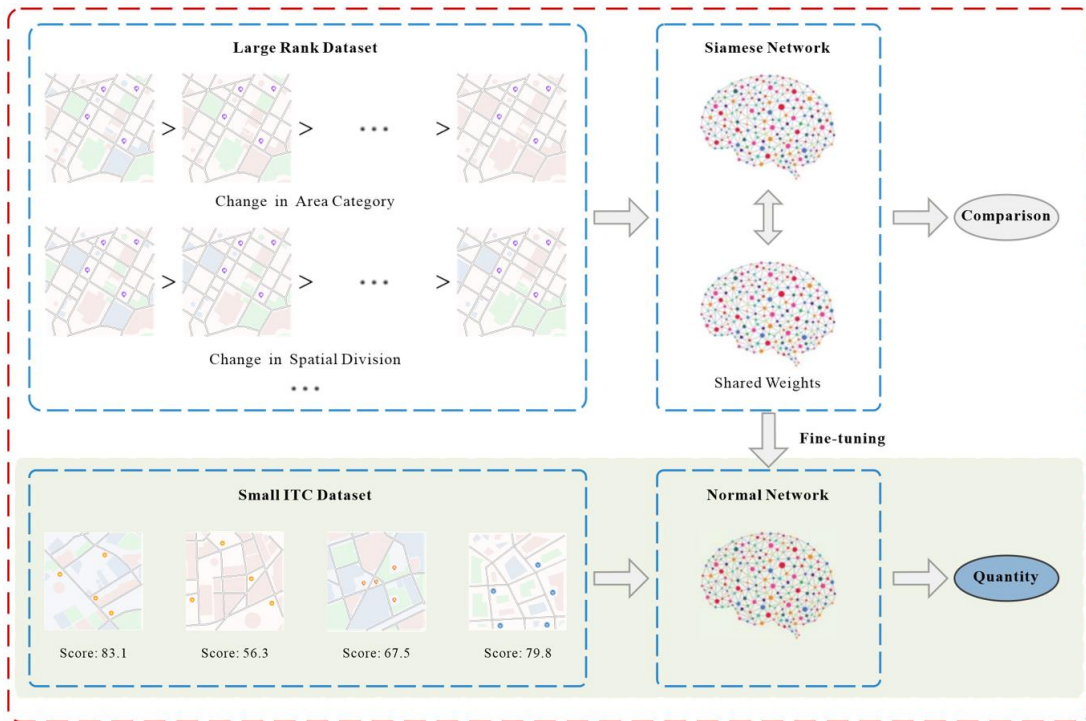


• 启示

- 人仅需观察几对地图和分数的对应关系便会打分

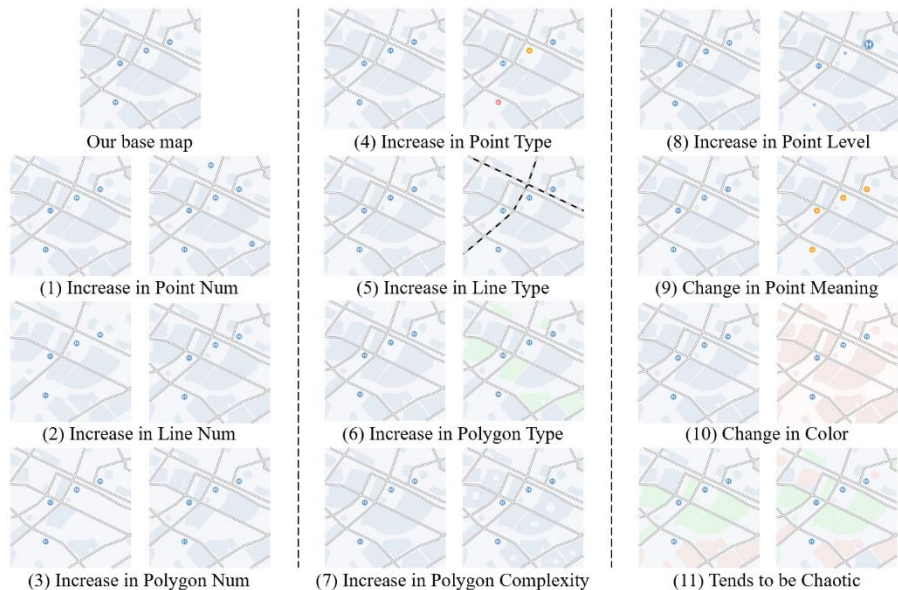
• 原因

- 人很容易感受地图的相对变化(有先验知识)



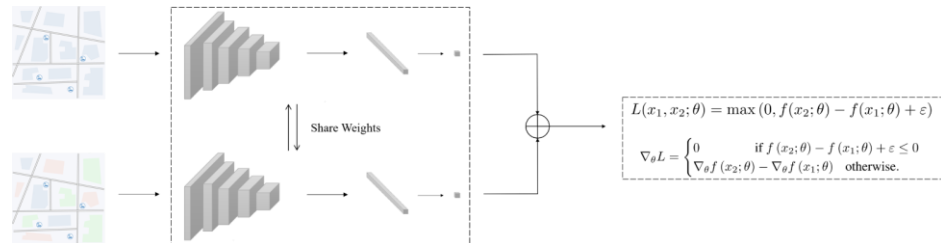
• 相对数值数据集

- 90张基图，进行11种不同方式的变换

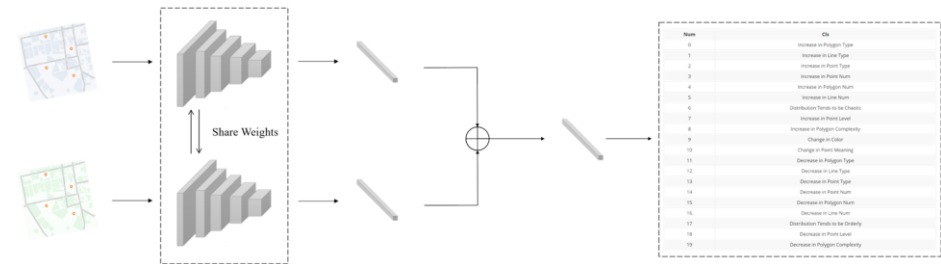


• 使用视觉模型预训练

- 基于排名(Rank)

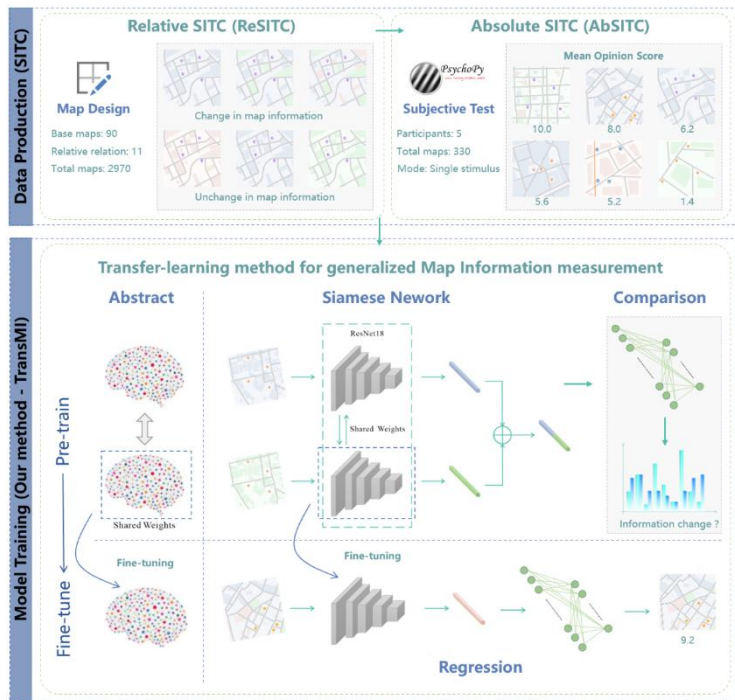


• 基于比较(Comparison)

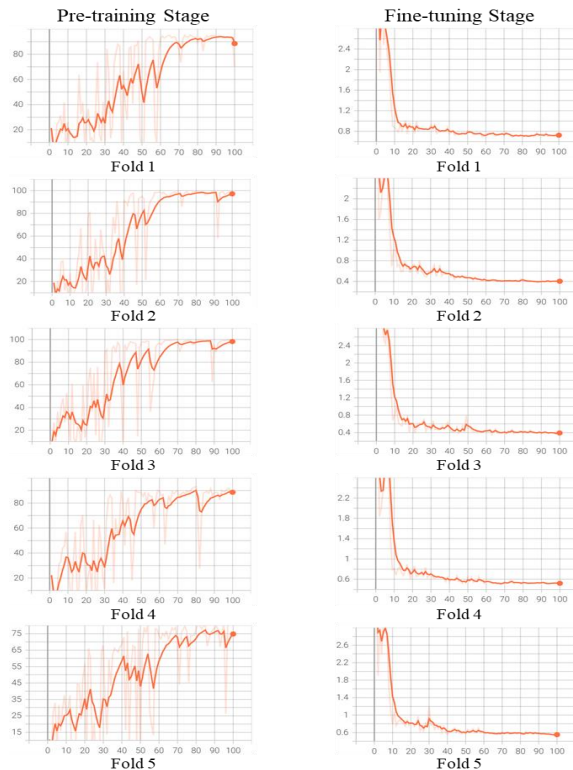


• 框架总览

- 预训练：孪生神经网络/相对数值数据集
- 微调：孪生神经网络一支/绝对数值数据集



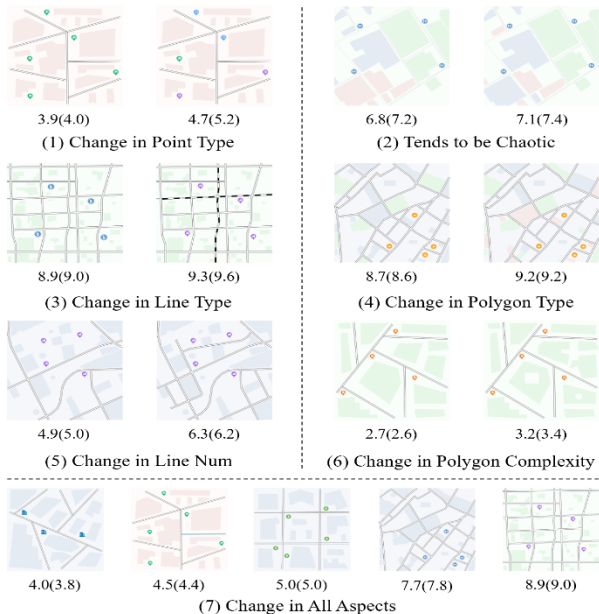
• 训练过程



预测结果

Model performance in the fine-tuning stage

Pretrained	K-fold Cross Validation (MAE)					Average
	1	2	3	4	5	
YES	0.385	0.692	0.352	0.496	0.537	0.492
NO	0.386	0.751	0.386	0.494	0.555	0.514



Notes that 3.9(4.0) represents that the ground truth is 4.0 and the prediction is 3.9

未来展望

- 大规模数据集构建(制图者截屏)
- 大规模预训练开展(ViT/Swin Transformer/...)
- 大模型泛化性能测试(OpenCLIP/BLIP/...)

模型

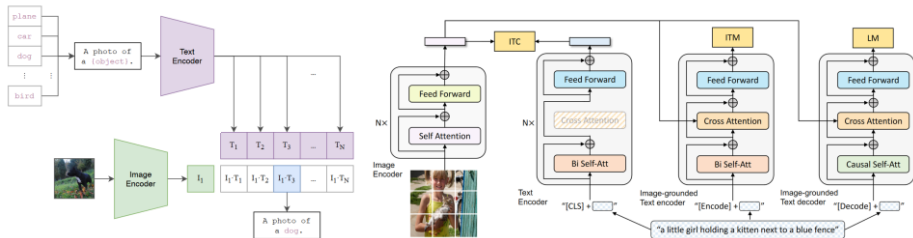
提示

CLIP

我们首先视觉编码器编码地图，文本编码器端编码构建的10种提示“这是一个广义信息量为[信息量数值1~10]的地图”，然后计算视觉特征与文本特征相似性并取最大值作为目标值。

BLIP

视觉编码器编码地图，我们使用提示“这个地图的广义信息量是多少呢？是1？2？3？还是...,10？”让模型生成答案。



OpenCLIP & BLIP



基于迁移学习的广义 地图信息量计算

汇报人：王成龙