Regression: Representation Space

Paper Reading by Yunyan Yan

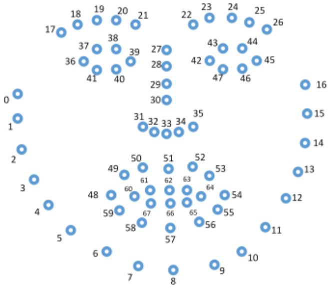
智能多媒体内容计算实验室

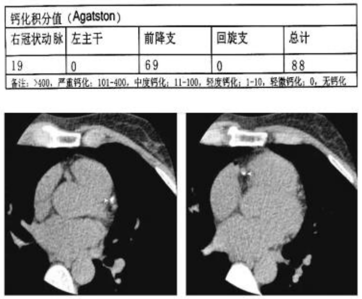
Intelligent Multimedia Content Computing Lab

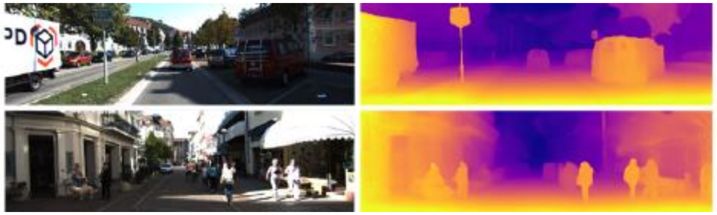
- Background
- Regression Metric Loss
- LARGE: Latent-Based Regression
- OrdinalCLIP
- Conclusion

Background

- Regression
  - 任务场景：
    - 关键点、年龄/骨龄、单目深度
    - (有序回归)病情等级、审美/评价
  - 现有方法：
    - L1 & L2
    - 特定先验(人脸关键点)
    - 转化为分类任务
  - 问题：
    - 高维特征表示+全连接 ---> **过拟合、难解释**
    - 特定先验 ---> **难迁移**
    - 分类 ---> **忽略序数关系**



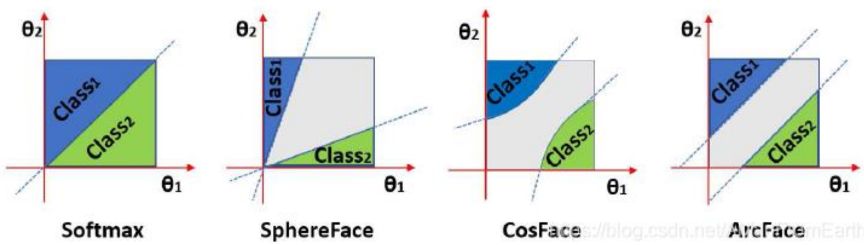




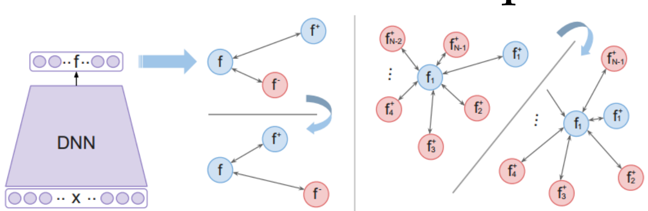
一个合适的低维特征空间

Background

- Metric Learning ---- 向量空间的映射
  - Proxy-based methods
    - 基于分类标签
    - L2-Softmax、CosFace、ArcFace



  - Pairwise-based methods
    - 基于样本对
    - Contrastive loss、Triplet loss、N-pair loss



Authors

Regression Metric Loss: Learning a Semantic Representation Space for Medical Images

Hanqing Chao, Jiajin Zhang, and Pingkun Yan(✉)

Department of Biomedical Engineering, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
{chaoh,zhangj41,yanp2}@rpi.edu

MICCAI 2022

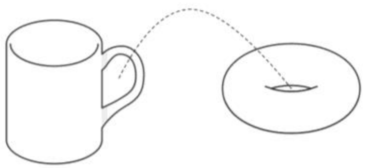
伦斯勒理工学院 生物医学工程

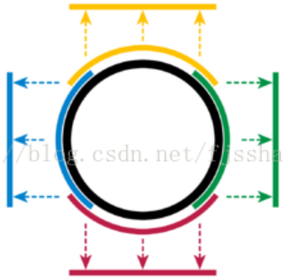
Motivation

- 现有问题
  - 高维特征缺乏可解释性
  - 空间内在结构不明
- 本文目的
  - 有语义的特征表示空间
  - 与标签空间等距的低维流形
- 具体实现
  - 表示空间的测地距离与标签空间的欧氏距离成正比
  - 根据流形上NN计算回归结果，代替全连接层

Methodology

- Riemannian Manifold
  - 微分流形+内积
  - 局部同胚于欧氏空间 $\mathbb{R}^n$





- Regression Metric Loss (RM-Loss)
  - 表示空间的测地距离 ∝ 标签空间的欧氏距离

$$D^o_{ij} = |s \times G(\boldsymbol{f}_i, \boldsymbol{f}_j) - ||\boldsymbol{y}_i - \boldsymbol{y}_j||_2|$$

  - 局部同胚 ---> 相同拓扑结构 ---> 测地距离≈欧氏距离

$$G(\boldsymbol{f}_i, \boldsymbol{f}_j), \boldsymbol{f}_j \in \mathcal{N}(\boldsymbol{f}_i) \quad ||\boldsymbol{f}_i - \boldsymbol{f}_j||_2$$

$$D_{ij} = |\ s \times ||\boldsymbol{f}_i - \boldsymbol{f}_j||_2 - ||\boldsymbol{y}_i - \boldsymbol{y}_j||_2\ |, \quad w_{ij} = \exp\left(-\frac{||\boldsymbol{y}_i - \boldsymbol{y}_j||_2^2}{2\sigma^2}\right) + \alpha$$

邻域大小　流形曲率

$$l' = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} D_{ij}}{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}},$$

Methodology

- Regression Metric Loss (RM-Loss)
  - 难样本挖掘(类似于ACSL)

$$m_{ij} = \begin{cases} 1, & \text{if } w_{ij} D_{ij} > \bar{l}^{(k)} \\ 0, & \text{if } w_{ij} D_{ij} <= \bar{l}^{(k)} \end{cases}, \quad \bar{l}^{(k)} = 0.9 \times \bar{l}^{(k-1)} + 0.1 \times \mathbb{E}^{(k)}(w_{ij} D_{ij})$$

$$\mathcal{L} = \frac{\sum_{p=1}^{N} \sum_{q=1}^{N} m_{ij} w_{ij} D_{ij}}{\sum_{p=1}^{N} \sum_{q=1}^{N} m_{ij} w_{ij}}$$

  - 推理阶段
    - 微分同胚：邻域 $\mathcal{N}^M(f_i) \subseteq M$ $\mathcal{N}^Y(y_i) \subseteq Y$ 一一映射
    - NN：测试样本的欧式邻域内，取训练样本加权

$$\hat{y}_t = \frac{\sum_{f_i \in \mathcal{N}_r(f_t)} a_i y_i}{\sum_{f_i \in \mathcal{N}_r(f_t)} a_i}, \quad a_i = \exp\left(-\frac{||\boldsymbol{f}_i - \boldsymbol{f}_t||_2^2}{2(r/3)^2}\right)$$

Experimental results

Performance comparison

Table 1. Regression performance and quality of the learned space on the full-size BAA and CAC datasets. The best results are in bold.

Method	Regression performance				Space quality			
	BAA		CAC		BAA 骨龄		CAC 冠脉钙化	
	MAE↓	R^2 ↑	MAE↓	R^2 ↑	D5↓	RV↓	D5↓	RV↓
MSE	7.021*	0.947	0.762*	0.930	8.872*	0.0690*	0.869*	0.1225*
L1	6.580*	0.952	0.668*	0.927	8.799*	0.0600	0.837*	0.1249*
OrdReg	7.061*	0.944	0.684*	0.937	9.368*	0.1450*	0.844*	0.1966*
ATrip+NN	6.799*	0.951	0.647*	0.939	9.022*	0.0617	0.834*	0.0874*
ATrip+L1	6.854*	0.949	0.660*	0.930	9.055*	0.0630*	0.855*	0.0813
RM-Loss (ours)	**6.438**	**0.954**	**0.617**	**0.944**	**8.614**	**0.0588**	**0.769**	**0.0806**

* $p < 0.05$ in the one-tailed paired t-test. The significant test was not applied on R^2.

度量损失+NN
效果更好

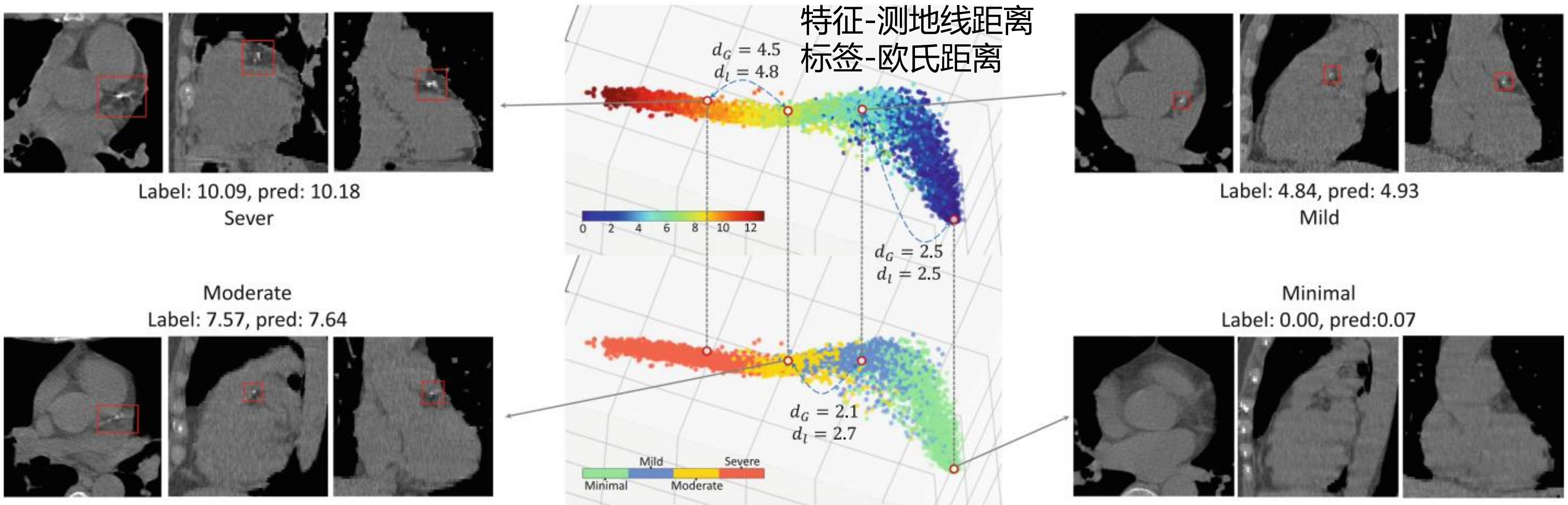
$$1 - \rho(G_M, D_Y)$$

评估流形M与标签Y是否等距

Backbone: Inception-v3

Experimental results

- Visualization on CAC



特征表示空间

Experimental results

□ Ablation Study

Table 2. Ablation study of σ, α, and m on BAA dataset. The best results are in bold.

过拟合 ------> 欠拟合 ------> 线性空间

邻域大小	σ	0.25	0.5	1.0	1.5	$+\infty$	0.5			
流形曲率	α	0.1					0.0	0.2	0.3	0.1
难例挖掘	m	w/								w/o
	MAE↓	6.555	**6.438**	6.642	6.759	6.823	6.496	6.591	6.741	6.520
	R^2 ↑	0.953	**0.954**	0.952	0.951	0.950	**0.954**	0.953	0.951	0.953
	D5↓	8.726	**8.614**	8.905	8.930	9.096	8.707	8.875	8.786	8.717
	RV↓	0.0614	**0.0588**	0.0658	0.0659	0.0699	0.0650	0.0614	0.0677	0.0641

- σ的取值非常接近临床诊断边界：骨龄(24 m; σ=0.5, 20.6 m)；CAC(10; σ=1.5, 8)
- α全局控制曲率

Experimental results

- Performance on 10%-size datasets

Table 3. The regression performance and the learned space quality on the 10%-size BAA and CAC dataset with the best results in bold.

Method	Regression performance				Space quality			
	10% BAA		10% CAC		10% BAA		10% CAC	
	MAE↓	R^2 ↑	MAE↓	R^2 ↑	D5↓	RV↓	D5↓	RV↓
MSE	8.721*	0.917	0.946*	0.895	11.204*	0.1054*	1.102*	0.1706*
L1	9.173*	0.906	0.875*	0.894	11.682*	0.1133*	1.028*	0.1529*
OrdReg	9.226*	0.908	0.849*	0.906	11.821*	0.2485*	1.010*	0.2189*
ATrip+NN	8.733*	0.911	0.861*	0.907	10.990	0.1018*	1.056*	0.1547*
ATrip+L1	9.017*	0.914	0.875*	0.908	11.208*	0.1016*	1.012*	0.1142
RM-Loss (ours)	**8.071**	**0.926**	**0.797**	**0.912**	**10.974**	**0.0878**	**0.971**	**0.1114**

* $p < 0.05$ in the one-tailed paired t-test. The significant test was not applied on R^2.

- Background
- Regression Metric Loss
- **LARGE: Latent-Based Regression**
  - Motivation
  - Methodology
  - Visualization
- OrdinalCLIP
- Conclusion

Authors

LARGE: Latent-Based Regression through GAN Semantics

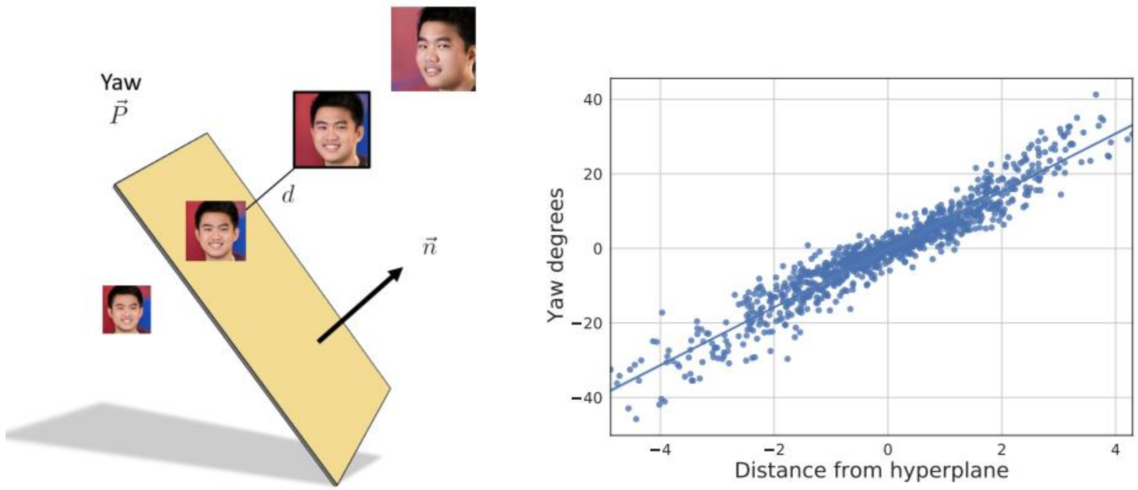
Yotam Nitzan* Tel-Aviv University
Rinon Gal* Tel-Aviv University
Ofir Brenner Tel-Aviv University
Daniel Cohen-Or Tel-Aviv University

CVPR 2022

以色列特拉维夫大学

Motivation

- 特征表达空间
  - GAN--图像编辑：语义解耦
  - GAN--潜在空间：成功的语义信息编码

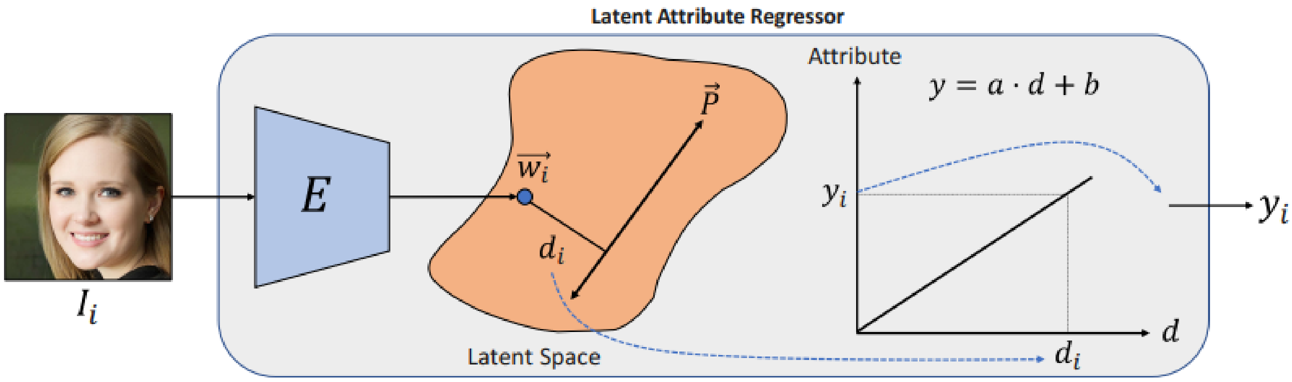


(a) (b)

- 本文思路
  - 微调预训练GAN模型，校准距离，沿属性方向回归

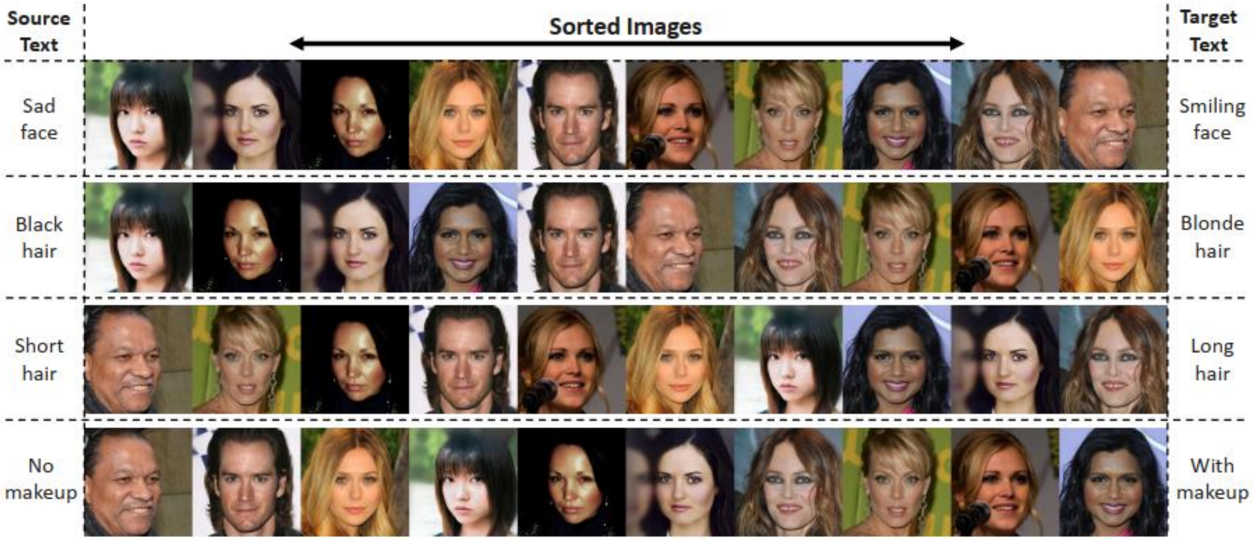
Methodology

Latent-based Regression



- 编码到超平面的距离
- StyleGAN的四个隐空间：高斯先验，GAN映射，每层分配，仿射变换 $\mathcal{Z}, \mathcal{W}, \mathcal{W}+$ and $\mathcal{S}$
- 不同层表达不同属性，对每层学习权重(不同层梯度平均幅度)

Visualization



- Background
- Regression Metric Loss
- LARGE: Latent-Based Regression
- **OrdinalCLIP**
  - Motivation
  - Methodology
- Conclusion

Authors

OrdinalCLIP: Learning Rank Prompts for Language-Guided Ordinal Regression

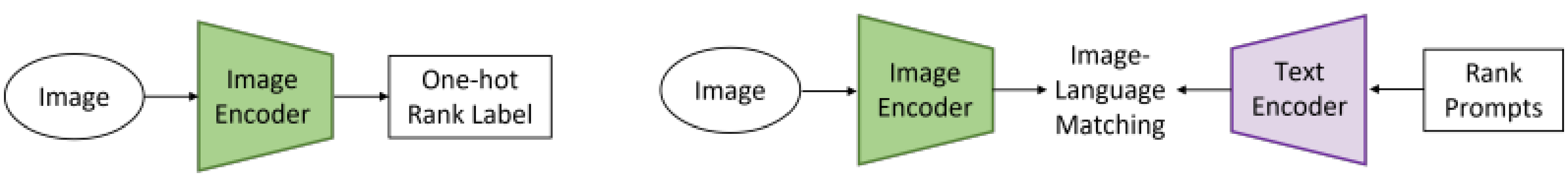
Wanhua Li[*,1], **Xiaoke Huang**[*,1], **Zheng Zhu**[2], **Yansong Tang**[1], **Xiu Li**[1], **Jie Zhou**[1], **Jiwen Lu**[†,1]
[1]Tsinghua University [2]PhiGent Robotics
wanhua016@gmail.com hxk21@mails.tsinghua.edu.cn
zhengzhu@ieee.org {tang.yansong, li.xiu}@sz.tsinghua.edu.cn
{jzhou, lujiwen}@tsinghua.edu.cn

NIPS 2022

清华大学 鲁继文

Motivation

- 任务
  - 有序分类
- 现有方法：
  - 连续 ---> 离散
  - 全连接层建模序数
- 本文思路
  - 语言先验学习分级概念
  - CLIP隐空间，图像-语言对齐，共同约束概念学习缓解过拟合

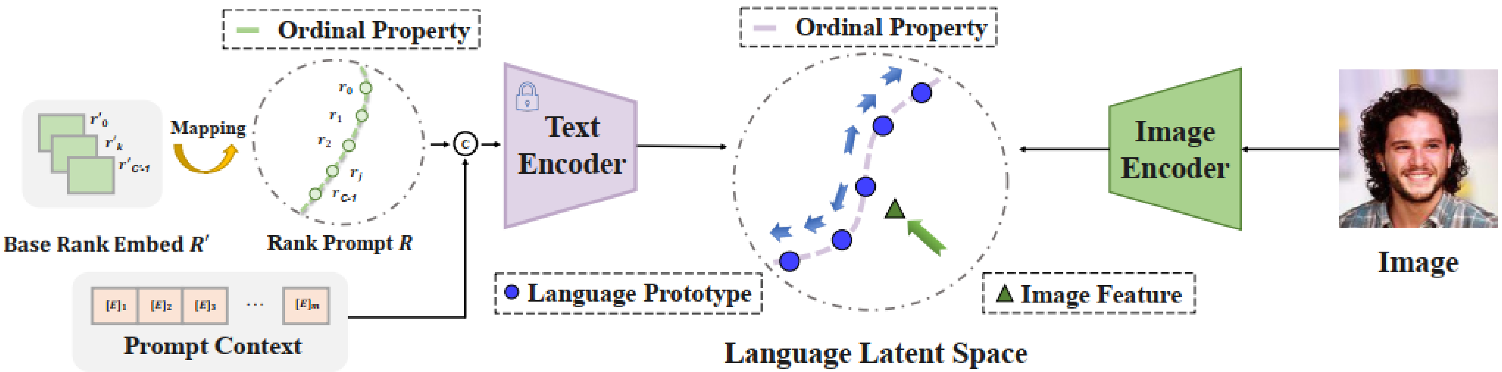


(a) Existing Methods for Ordinal Regression

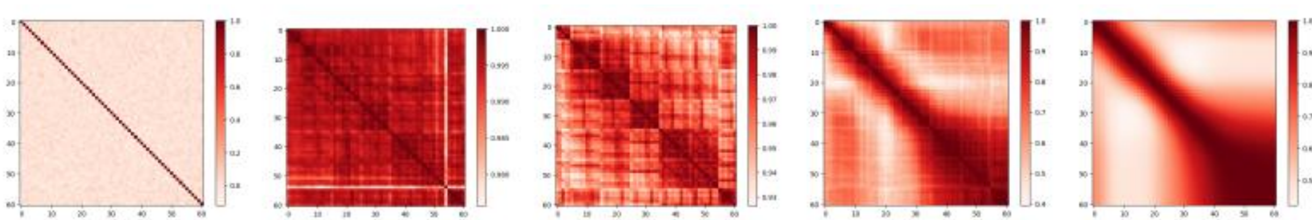
(b) Language Powered Paradigm for Ordinal Regression (Ours)

Methodology

OrdinalCLIP



- 语言模型固定(推理时无需使用)，图像编码可学习
- 内积计算相似度，KL散度约束接近单位阵
- 上下文可学习，序数插值嵌入
- CLIP直接用于序数回归的效果较差(人工prompt序数关系差)



(a) CNN Baseline (b) CLIP, Rand. Ctx (c) CLIP, Init. Ctx (d) CoOp (e) OrdinalCLIP

- Background
- Regression Metric Loss
- LARGE: Latent-Based Regression
- OrdinalCLIP
- **Conclusion**

Conclusion

- 特征表示空间
  - 度量学习约束
  - 具有优秀性质的预训练模型隐空间
    - GAN
    - CLIP

Thanks for Attention!