

Taskonomy: Disentangling Task Transfer Learning

Amir R. Zamir^{1,2} Alexander Sax^{1*} William Shen^{1*} Leonidas Guibas¹ Jitendra Malik² Silvio Savarese¹
¹ Stanford University ² University of California, Berkeley Apple Inc

<https://arxiv.org/pdf/1804.08328.pdf>

CVPR 2018, Best Paper Award

Summary

By Yohei Kawakami
2019/01/16



Abstract

転移学習の転移の容易さを類似度で評価する

26のヴィジュアルタスクごとに転移学習させあって相性をグラフ化
各転移学習モデルごと + 転移学習させずに1からスクラッチ を比較!

What is this thesis for?

転移学習の転移の容易さを26の
ビジュアルタスクごとに評価した

Where is an important point compared to previous researches?

転移学習の転移の容易さが分かれば
最適な転移学習モデルを選択でき教師ラベルの
データを減らしても精度上げられる

Where are the key points of technology and method?

実行可能なすべてのソースとターゲット
の組み合わせ(約3000)についてtransfer functionを
トレーニングし、AHP(Analytic Hierarchy Process)
を使い、タスク類似度行列を作成

How to verified whether it is valid?

$Wt' = Wt / Wt.T$ のように割り、行列において
タスク s_i から学習したモデルが
タスク s_j から学習したモデルと比較して
どのくらい優れているかで示す

Is there discussions?

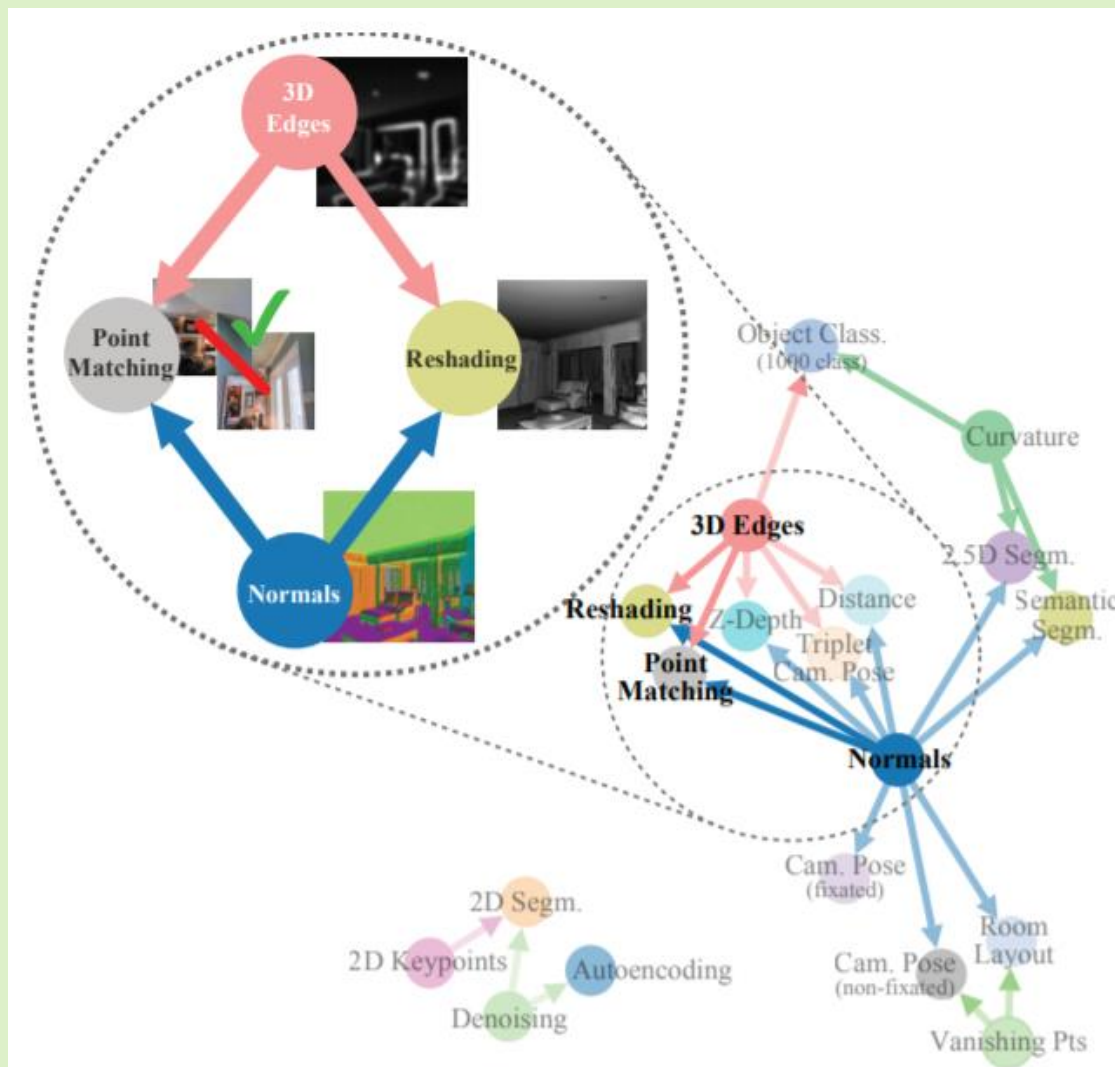
様々なアーキテクチャおよびデータ
セットに対して検証したが、原則的に
モデルおよびデータ特有のものである
ので一般化等のさらなる検討が必要

Which reserches should I read next?

Transitive invariance for self-supervised visual
representation learning.

Summary:

Taskonomy(タスク間の親和性の分類)



問題提起

深層学習にかかるコスト、計算リソース、または時間は制限される

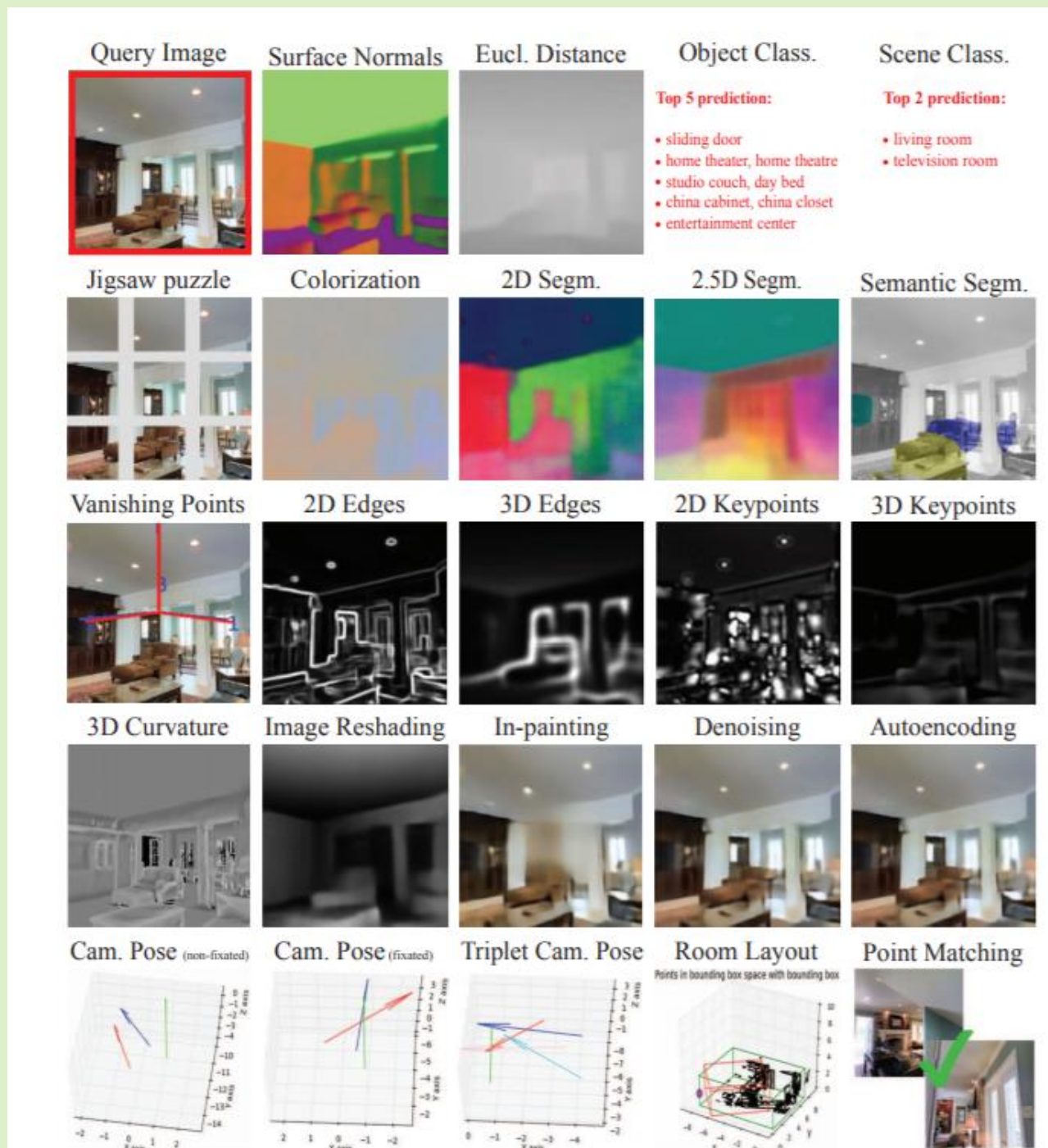
限られた資源の中でタスクのパフォーマンスを最大化させたい!

$$\text{Taskonomy} = \text{task} \times \text{taxonomy}$$

26のビジュアルタスクで実行可能な転移学習の組み合わせをすべて行い関連の深い(転移学習させやすい)組み合わせを見つけ出す手法を提案)

Summary:

Dataset



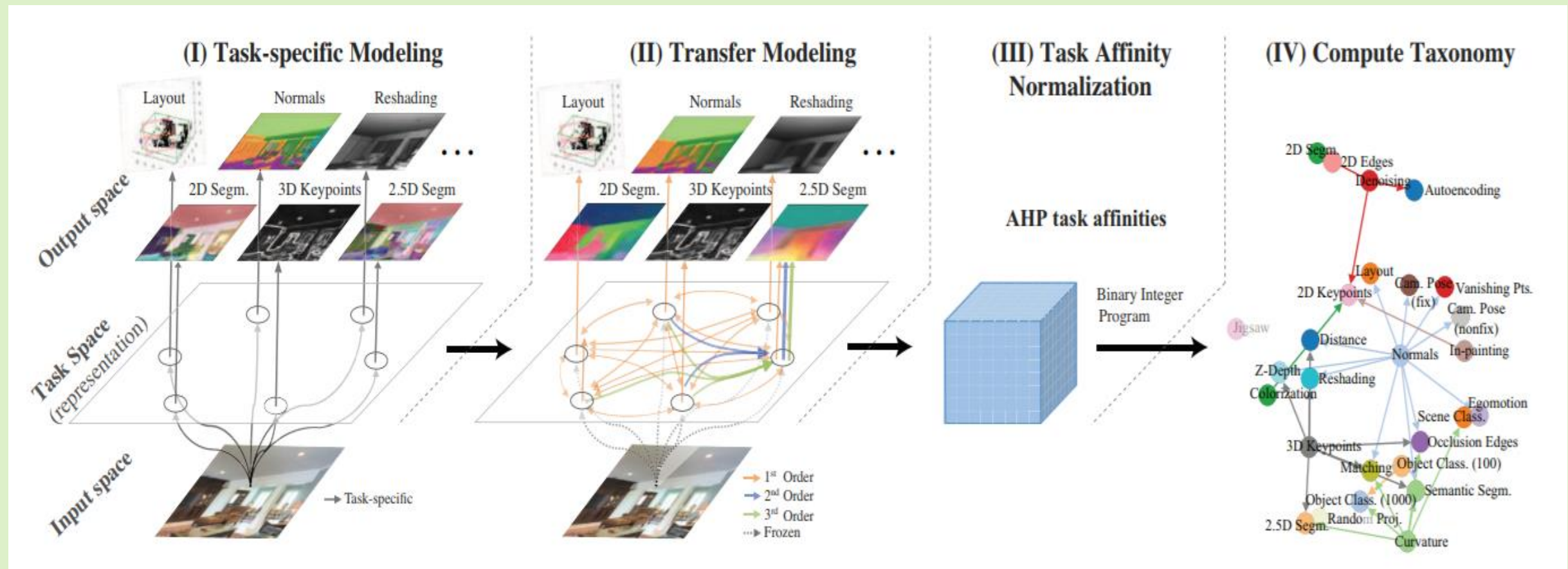
dataset

約600棟の建物から400万枚の
屋内シーンの画像データセッ
トを作成

すべての画像には、すべての
タスクのアノテーションラベ
ルあり

Summary:

Method



1 タスク固有モデルの構築
転移学習させる
元になるモデル構築
のための教師あり学習を
行いタスク固有モデルを学習

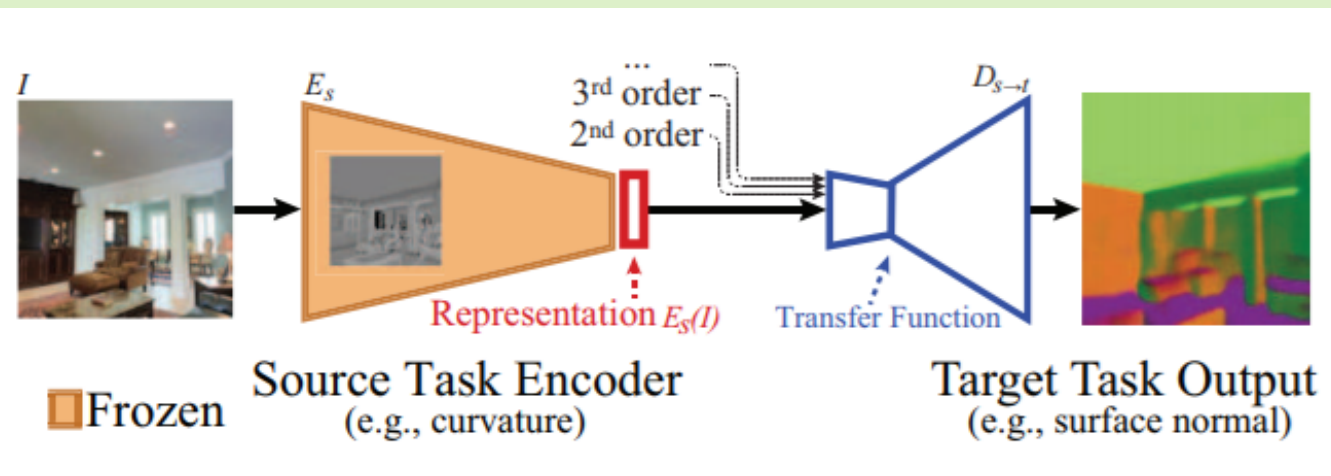
2 転移学習
1で学習した固有モデルを
使い別のタスクを転移学習
全ての転移可能な組み合わせ
の転移学習を行う

3 正規化
階層分析法 (AHP) を
使用しテーブル化後
正規化する

4 ハイパーグラフを作成
転移タスクの最適化と
すべての転移方式の
パフォーマンスを予測し、
ハイパーグラフを作成

Summary:

Method



転移学習モデル

学習したSource TaskをEncoderに入力しTransfer Functionを学習してTarget Task をOutputする

Transfer Funcion

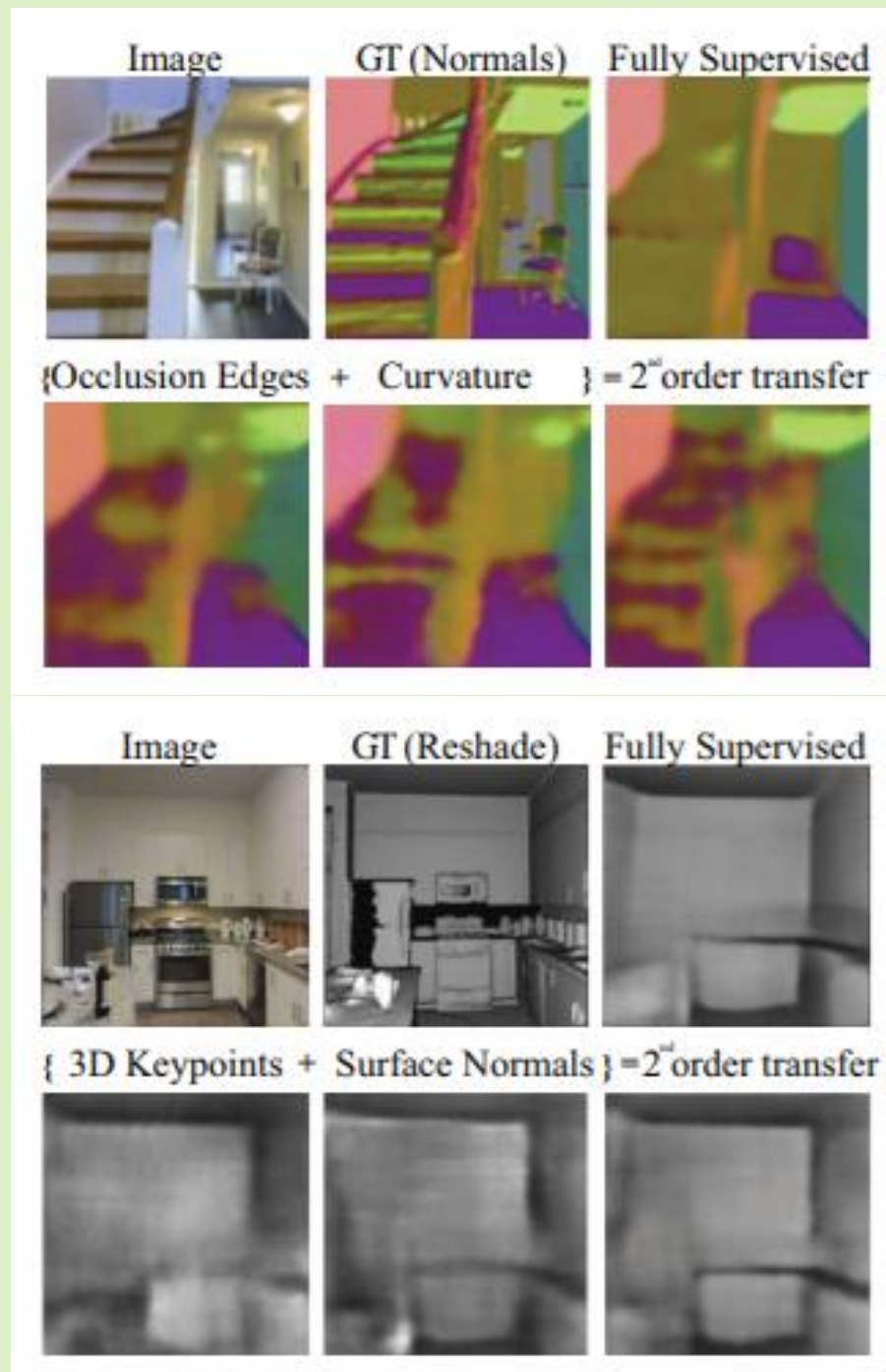
$f_t(I)$: Image I に対する t の ground truth
 $E_s(I)$: source s の encoder から抽出した画像 I の Representation
 L_t : Loss Function
 L_t を最小にする θ で parameter化

$$D_{s \rightarrow t} := \arg \min_{\theta} \mathbb{E}_{I \in \mathcal{D}} \left[L_t \left(D_{\theta}(E_s(I)), f_t(I) \right) \right]$$

$D_{s \rightarrow t}$ のパフォーマンスは、タスクの親和性として有用な測定基準

Summary:

Higher-Order-Transfers



高次元の転移

Higher-Order Transfers

補足的な情報も含まれる
 良い結果が得られる可能性が
 低い組み合わせは除外

Summary:

AHP



Intermediate pairwise matrix W_t

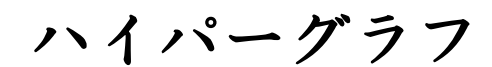
正規化する前(左)と後(右)
の一次タスク親和性行列
(AHP)

Matrix element

$W_t/W_t.T$ とすることで
 W'_{ij} (要素)
 S_i が S_j より良好となった
勝率

$$w'_{i,j} = \frac{\mathbb{E}_{I \in \mathcal{D}_{test}} [D_{s_i \rightarrow t}(I) > D_{s_j \rightarrow t}(I)]}{\mathbb{E}_{I \in \mathcal{D}_{test}} [D_{s_i \rightarrow t}(I) < D_{s_j \rightarrow t}(I)]}$$

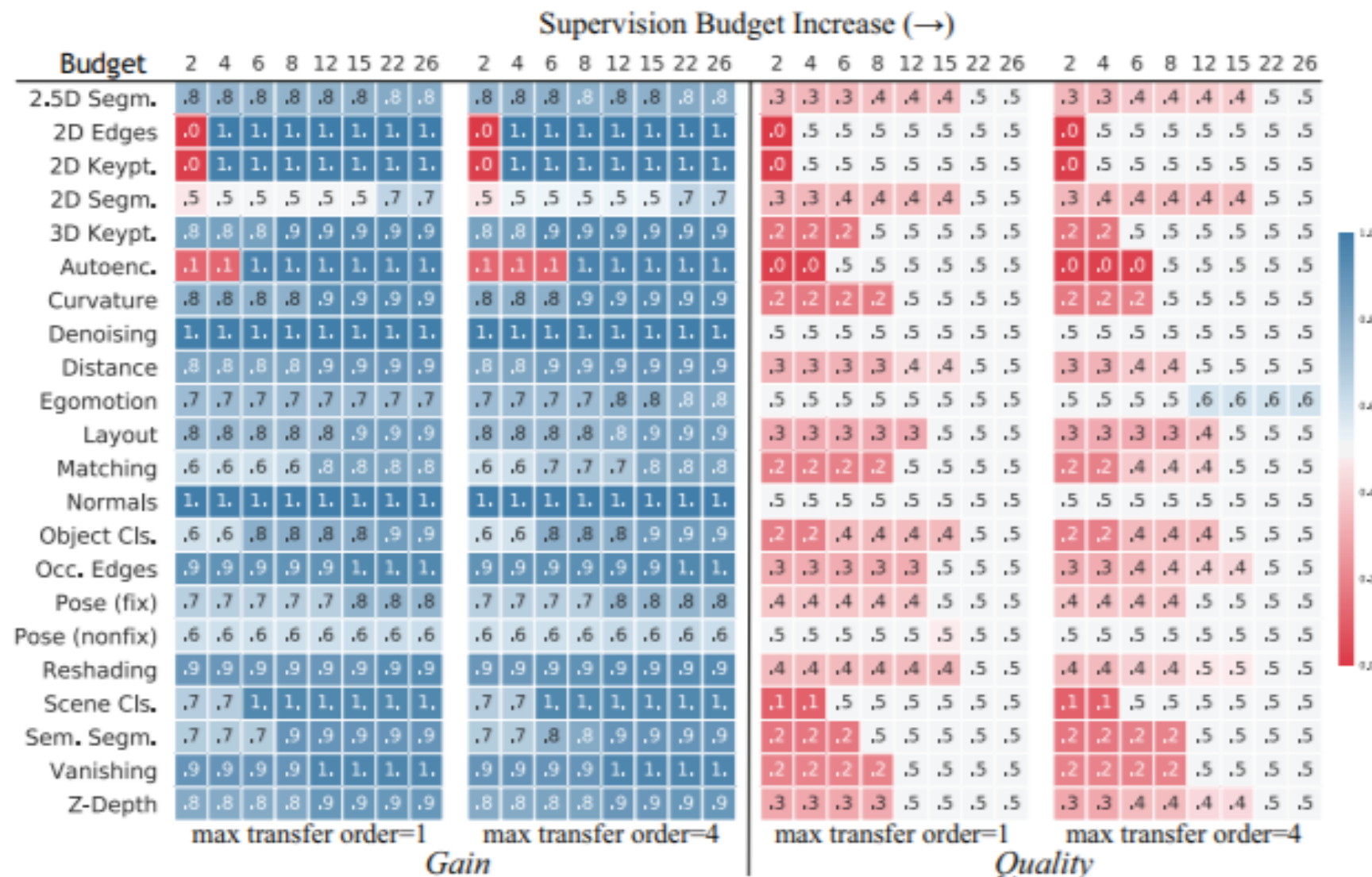
Taskonomy



X軸が監督予算
Y軸がソースの数

Summary:

Supervision Budget Increase



Gain

転移ネットワークと同じ数のトレーニングデータを使用して最初からトレーニングしたネットワークに対する勝率 (%)

Quarity:

(ゴールドスタンダード) な 120kの画像でネットワークに対する勝率(%)

Summary:

Generalization to Novel Task

Order Increase (→)								
Order	1	2	3	4	1	2	3	4
2.5D Segm.	.1	.1	.1	.9	.0	.0	.0	.0
2D Edges	.8	.8	.8	.8	.4	.4	.4	.4
2D Keypt.	.8	.8	.9	.9	.2	.2	.2	.2
2D Segm.	.8	.9	.9	.9	.2	.3	.3	.3
3D Keypt.	.0	.1	.9	.9	.0	.0	.0	.0
Autoenc.	.9	1.	1.	1.	.0	.0	.0	.0
Curvature	.9	.9	.9	.9	.3	.4	.4	.4
Denoising	.7	.8	.8	.8	.5	.6	.6	.6
Distance	.7	.7	.7	.7	.4	.4	.4	.4
Egomotion	.0	.9	.9	.9	.0	.0	.0	.0
Layout	.9	.9	.9	.9	.1	.2	.2	.2
Matching	.6	.6	.6	.6	.5	.5	.5	.5
Normals	.6	.7	.7	.7	.2	.4	.4	.5
Object Cls.	.9	.9	.9	.9	.4	.5	.5	.5
Occ. Edges	.9	.9	.9	.9	.4	.4	.4	.4
Pose (fix)	.9	.9	.9	.9	.4	.5	.5	.5
Pose (nonfix)	.9	1.	1.	1.	.0	.0	.1	.1
Reshading	.8	.8	.8	.8	.3	.3	.4	.4
Scene Cls.	.8	.8	.8	.8	.4	.4	.4	.4
Sem. Segm.	.5	.5	.5	.5	.4	.4	.4	.4
Vanishing	.7	.8	.8	.8	.2	.2	.2	.2
Z-Depth	.9	.9	.9	.9	.2	.2	.2	.2

Task	scratch	ImageNet[49]	Wang.[93]	Agrawal.[1]	Zamir.[97]	Zhang.[100]	Noroozi.[65]	full sup.	Taxonomy
Depth	88	88	93	89	88	84	86	43	-
	.03	.04	.04	.03	.04	.03	.03	.02	.02
Scene Cls.	80	52	83	74	74	71	75	15	-
	3.30	2.76	3.56	3.15	3.17	3.09	3.19	2.23	2.63
Sem. Segm.	78	79	82	85	76	78	84	21	-
	1.74	1.88	1.92	1.80	1.85	1.74	1.71	1.42	1.53
Object Cls.	79	54	82	76	75	76	76	34	-
	4.08	3.57	4.27	3.99	3.98	4.00	3.97	3.26	3.46
Normals	97	98	98	98	98	97	97	6	-
	.22	.30	.34	.28	.28	.23	.24	.12	.15
2.5D Segm.	80	93	92	89	90	84	87	40	-
	.21	.34	.34	.26	.29	.22	.24	.16	.17
Occ. Edges	93	96	95	93	94	93	94	42	-
	.16	.19	.18	.17	.18	.16	.17	.12	.13
Curvature	88	94	89	85	88	92	88	29	-
	.25	.28	.26	.25	.26	.26	.25	.21	.22
Egomotion	79	78	83	77	76	74	71	59	-
	8.60	8.58	9.26	8.41	8.34	8.15	7.94	7.32	6.85
Layout	80	76	85	79	77	78	70	36	-
	.66	.66	.85	.65	.65	.62	.54	.37	.41

新しいタスクに対する
一般化

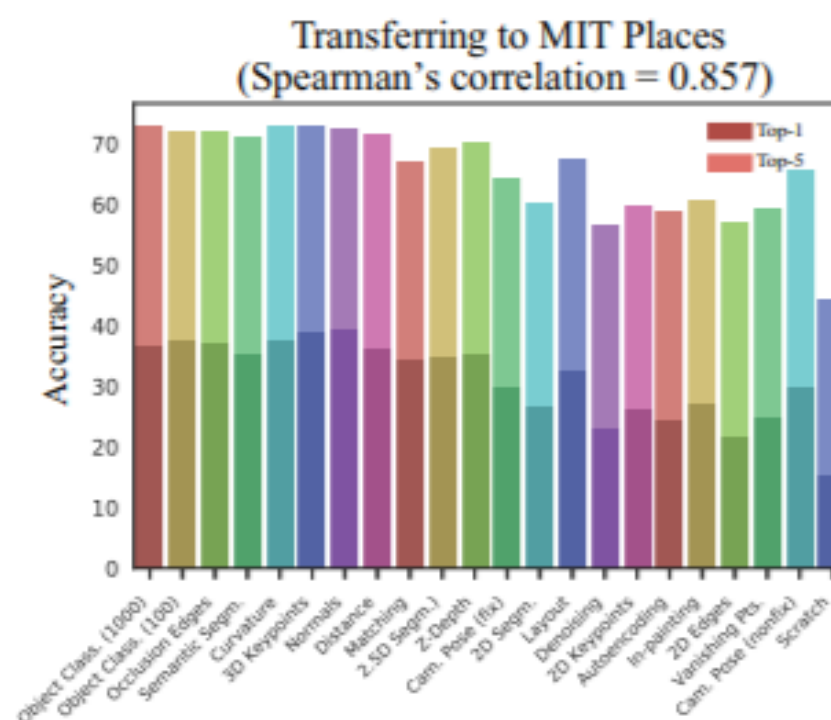
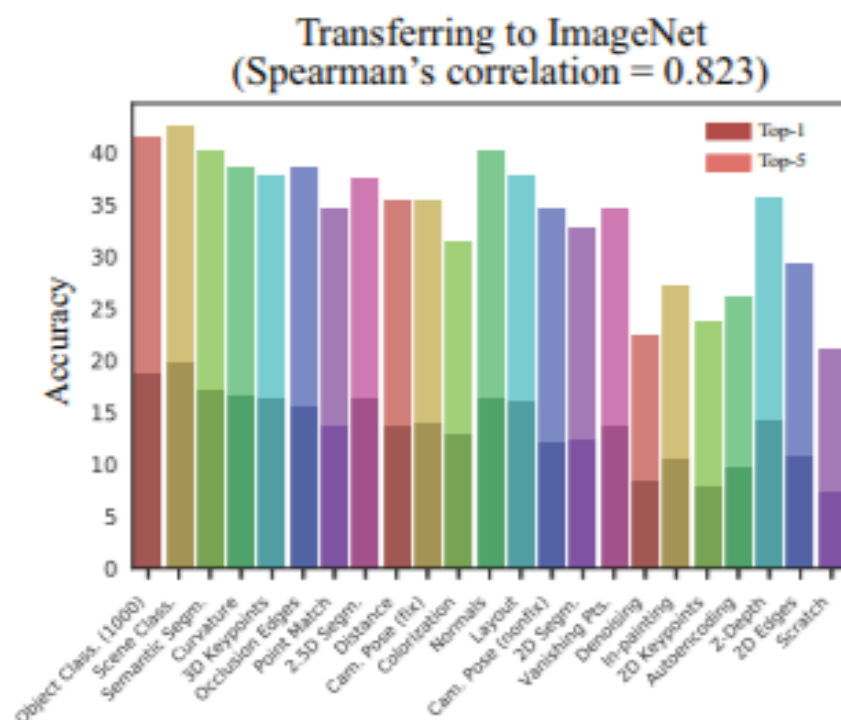
Taxonomyにより効率的な
転移学習を行うことで
既存の手法よりパフォー
マンスが向上した

Summary:

Structure Significance and Evaluating the discovered structure on other datasets:



適切に転移学習を選んだ
(green)場合とランダムに
選んだ(gray)場合の比較

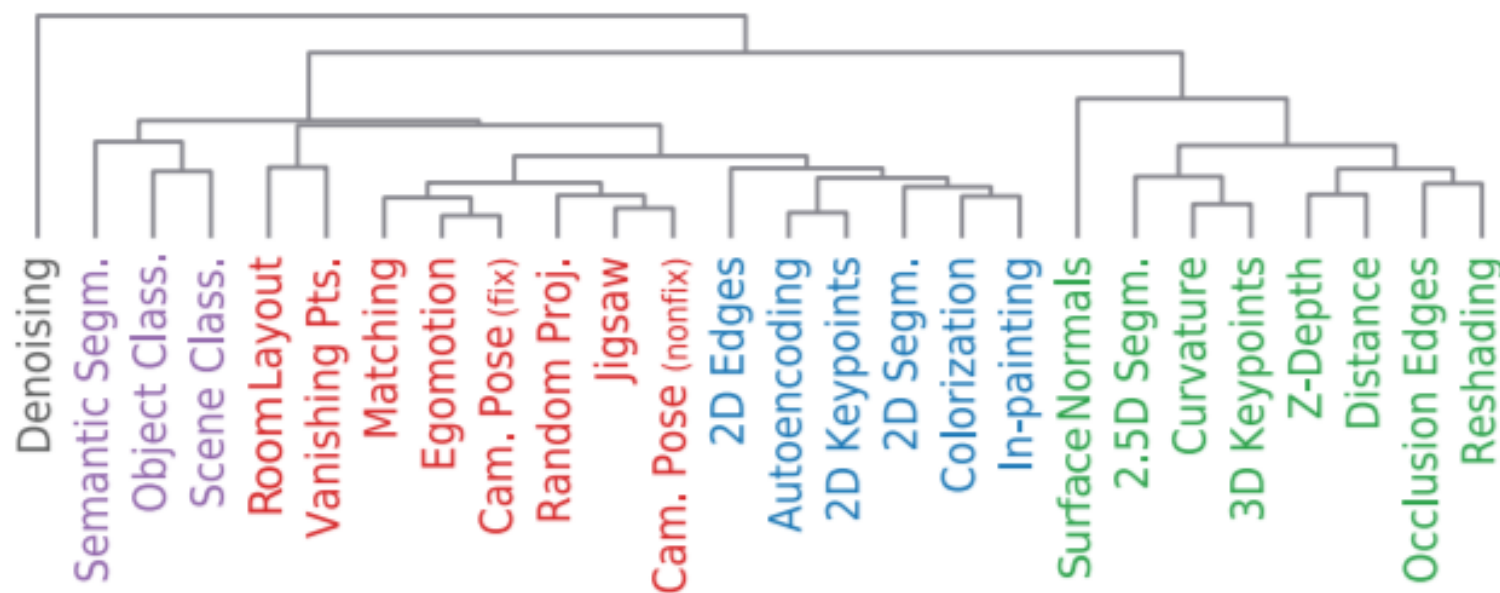


今回のデータセットは
MITのデータセットと相
関がありTaxonomyの手法
は他のデータセットにも
応用できそう

Summary:

Task Similarity Treer

Task Similarity Tree Based on Transferring-Out



タスク類似度ツリー

タスクの特徴ベクトル
として正規化された
親和性行列 P の列を使用して、
タスクの凝集クラスタリングから
得られる

Abstract

転移学習の転移の容易さを類似度で評価する

26のヴィジュアルタスクごとに転移学習させあって相性をグラフ化
各転移学習モデルごと + 転移学習させずに1からスクラッチ を比較!

What is this thesis for?

転移学習の転移の容易さを26の
ビジュアルタスクごとに評価した

Where is an important point compared to previous researches?

転移学習の転移の容易さが分かれば
最適な転移学習モデルを選択でき教師ラベルの
データを減らしても精度上げられる

Where are the key points of technology and method?

実行可能なすべてのソースとターゲット
の組み合わせ(約3000)についてtransfer functionを
トレーニングし、AHP(Analytic Hierarchy Process)
を使い、タスク類似度行列を作成

How to verified whether it is valid?

$Wt' = Wt / Wt.T$ のように割り、行列において
タスク s_i から学習したモデルが
タスク s_j から学習したモデルと比較して
どのくらい優れているかで示す

Is there discussions?

様々なアーキテクチャおよびデータ
セットに対して検証したが、原則的に
モデルおよびデータ特有のものである
ので一般化等のさらなる検討が必要

Which reserches should I read next?

Transitive invariance for self-supervised visual
representation learning.