



画像変換ネットワークによる 連続学習

松本 農人

電気通信大学 総合情報学科

柳井研究室

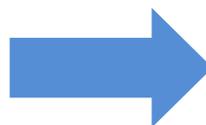
連続学習とは

- 知識の保持 + 新たな知識の学習



午前 自転車の練習

既に学習した知識



午後 足し算の勉強

新たに学習した知識

連続学習とは

- 知識の保持 + 新たな知識の学習



午前 自転車の練習

既に学習した知識



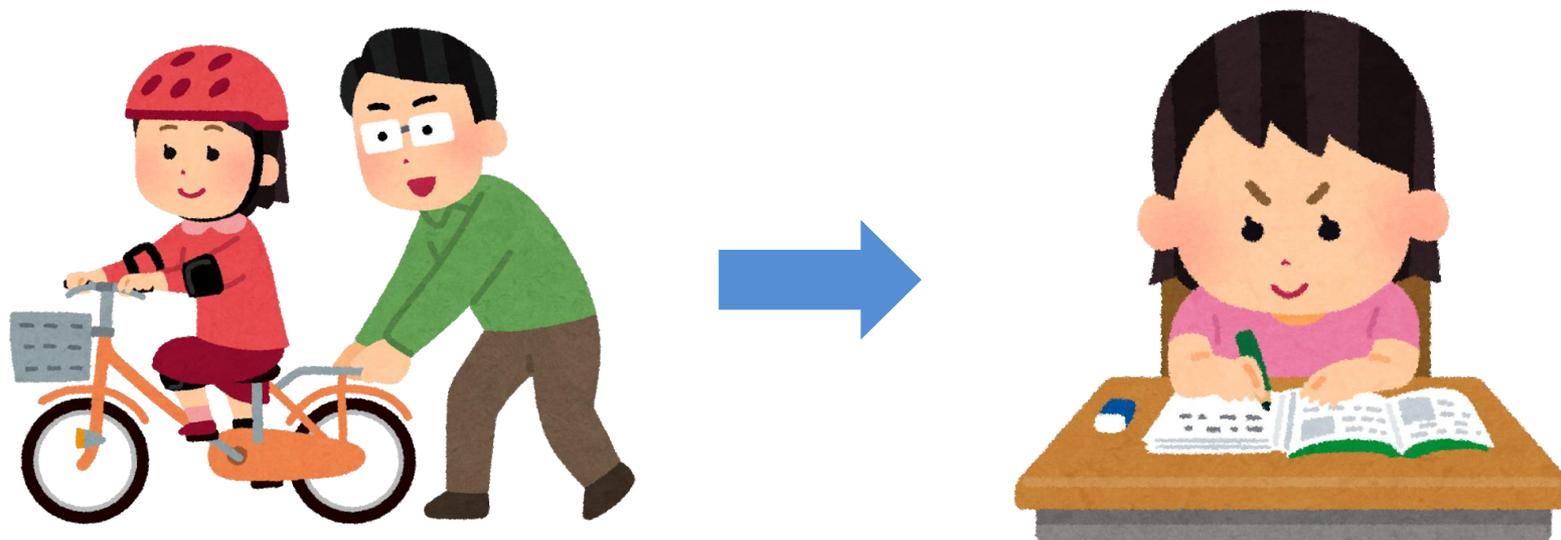
午後 足し算の勉強

新たに学習した知識

忘れない

連続学習とは

- 知識の保持 + 新たな知識の学習



昔学習した知識を忘れずに
新たに異なるタスクを学習

連続学習の特徴

- 1つのシステムで様々なタスクに対応



汎用人工知能



デバイス実装

CNNでの連続学習

- **問題点** 昔の知識を忘れてしまう
 - まず動物の分類問題を学習

タスク1



動物

馬

CNNでの連続学習

- **問題点** 昔の知識を忘れてしまう
 - ・ 次に花の分類問題を学習

タスク2



花 after 動物 →

パンジー

CNNでの連続学習

● 問題点 昔の知識を忘れてしまう

- 「花 after 動物」で動物の分類問題をやると...

タスク1



タスク2



花 after 動物 →

鹿？

性能劣化

パンジー

パラメータ上書き

CNNでの連続学習

● 問題点 昔の知識を忘れてしまう

- 「花 after 動物」で動物の分類問題をやると...



破壊的忘却を回避する手法

- **どれも分類タスクのみ**

- EWC [Kirkpatrick et al. 2016]
- LwF [Li and Hoiem 2016]
- Piggyback [Mallya et al. ECCV 2018]

- **画像変換タスクでは行われていない**

画像変換ネットワークの連続学習

- “一つのモデル”で“異なる”
画像変換タスクの連続学習

領域分割



Piggyback



スタイル変換



着色



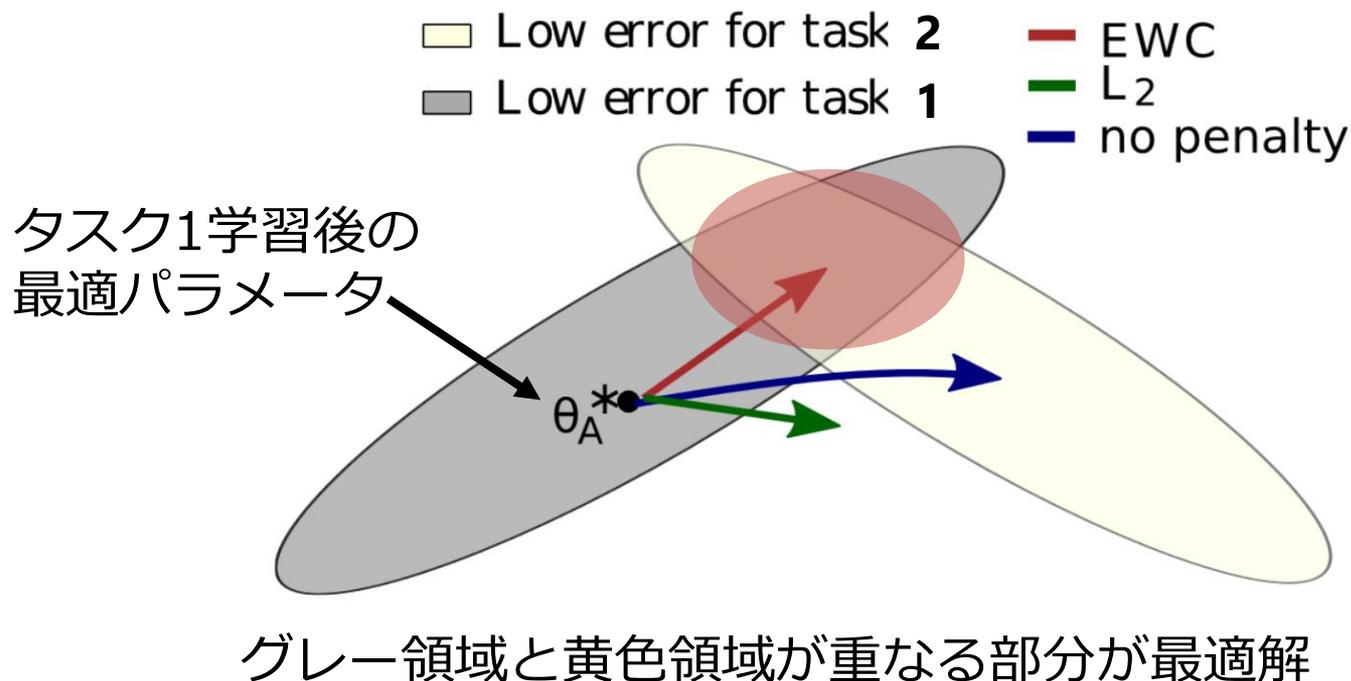
input



output

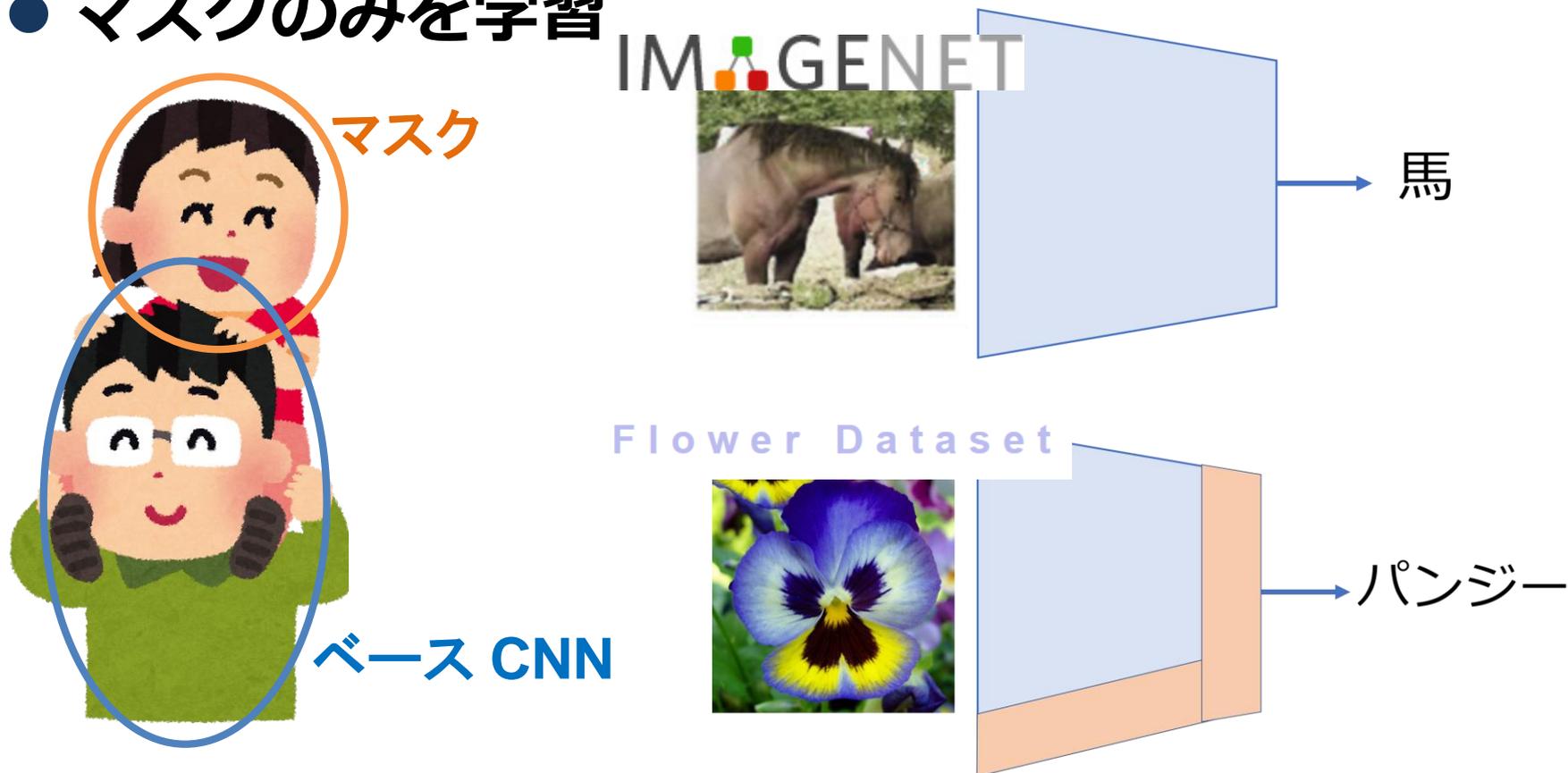
Elastic Weight Consolidation (EWC)

- タスク1とタスク2の両方の最適解を求める
- 問題: 最適領域が重ならない場合



Piggyback

- タスク毎に重要な重みをマスクで選択
- マスクのみを学習

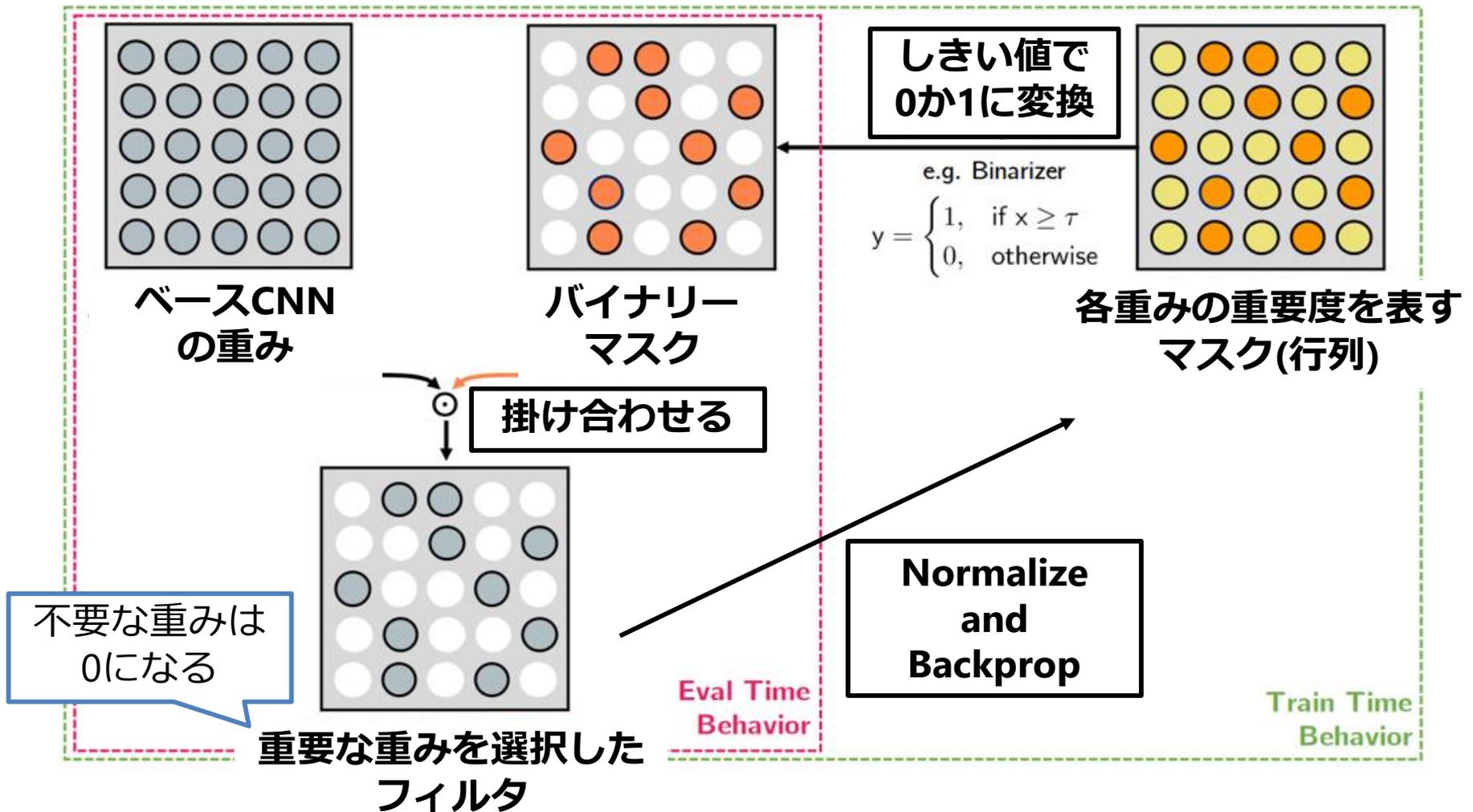


Encoder-Decoder CNN + Piggyback

- ベースタスク1
- 追加タスク2
- 追加タスク3
- 追加タスク4



マスクの学習

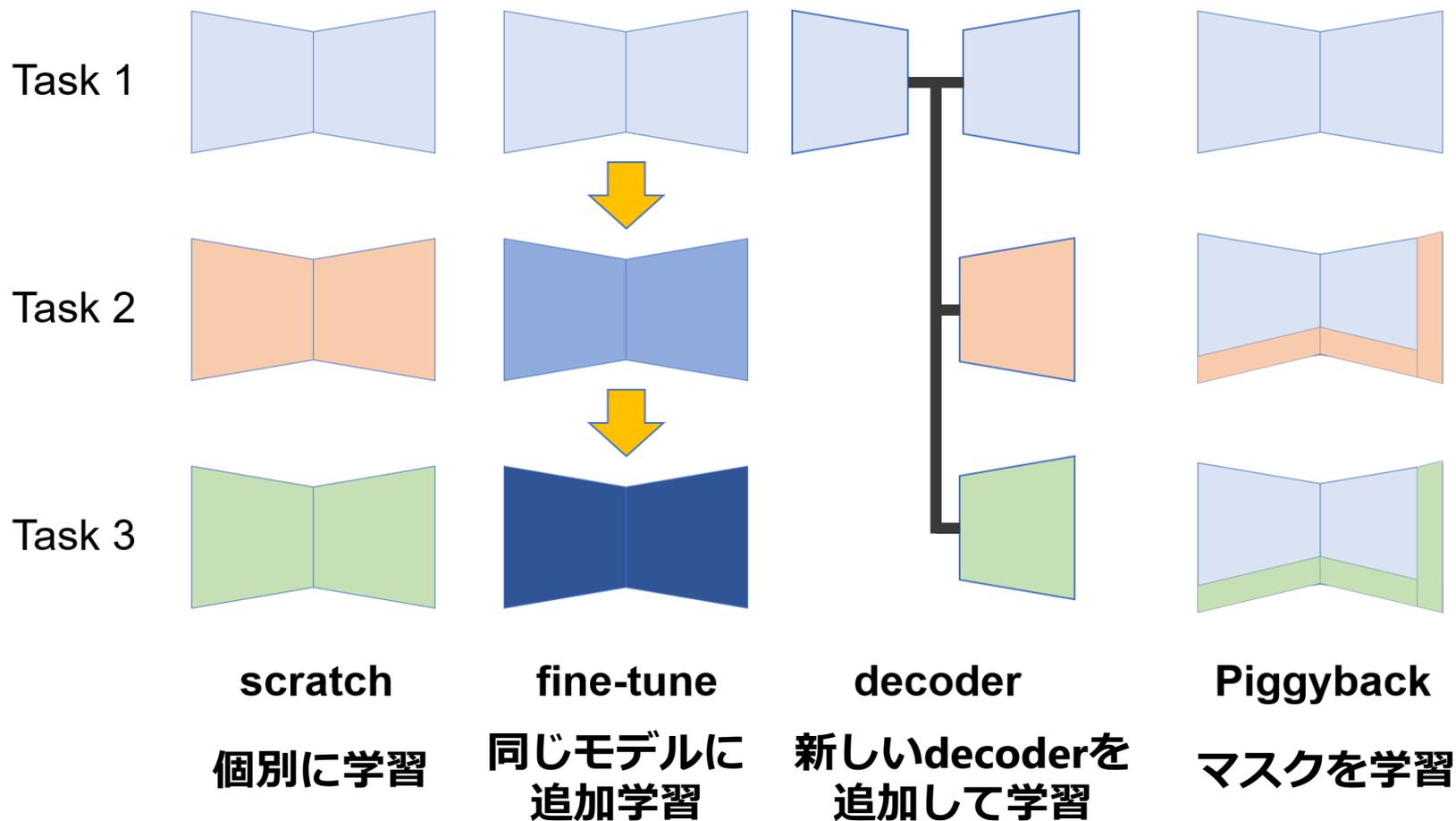


タスクと内容

タスク番号	データセット	内容	評価
タスク 1	MS COCO	領域分割	mIoU(%)
タスク 2	Pascal VOC	領域分割	mIoU(%)
タスク 3	MS COCO	濃淡画像着色	SSIM
タスク 4	MS COCO	スタイル変換 (Gogh)	SSIM
タスク 5	MS COCO	スタイル変換 (Munk)	total loss

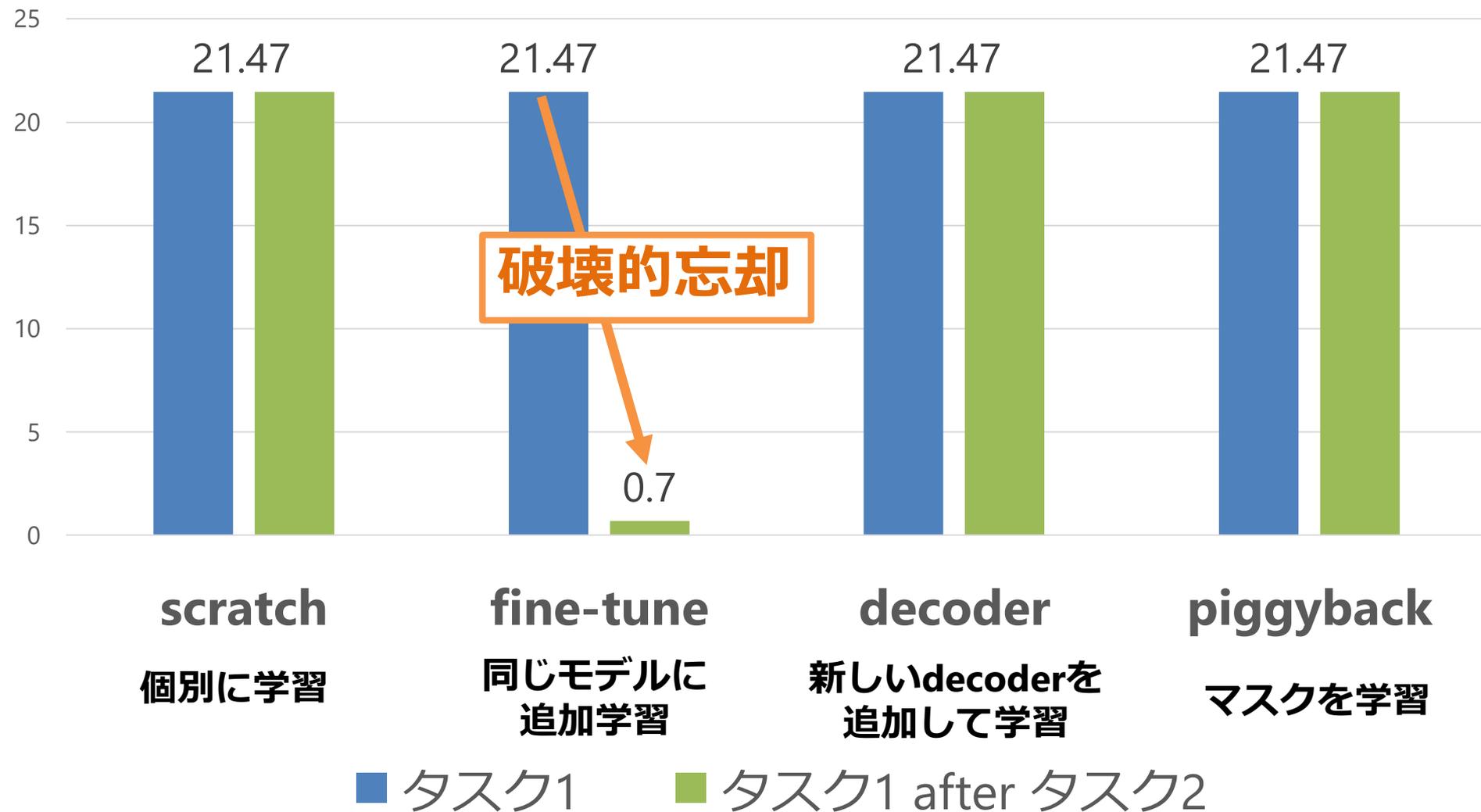


ベースラインと提案手法



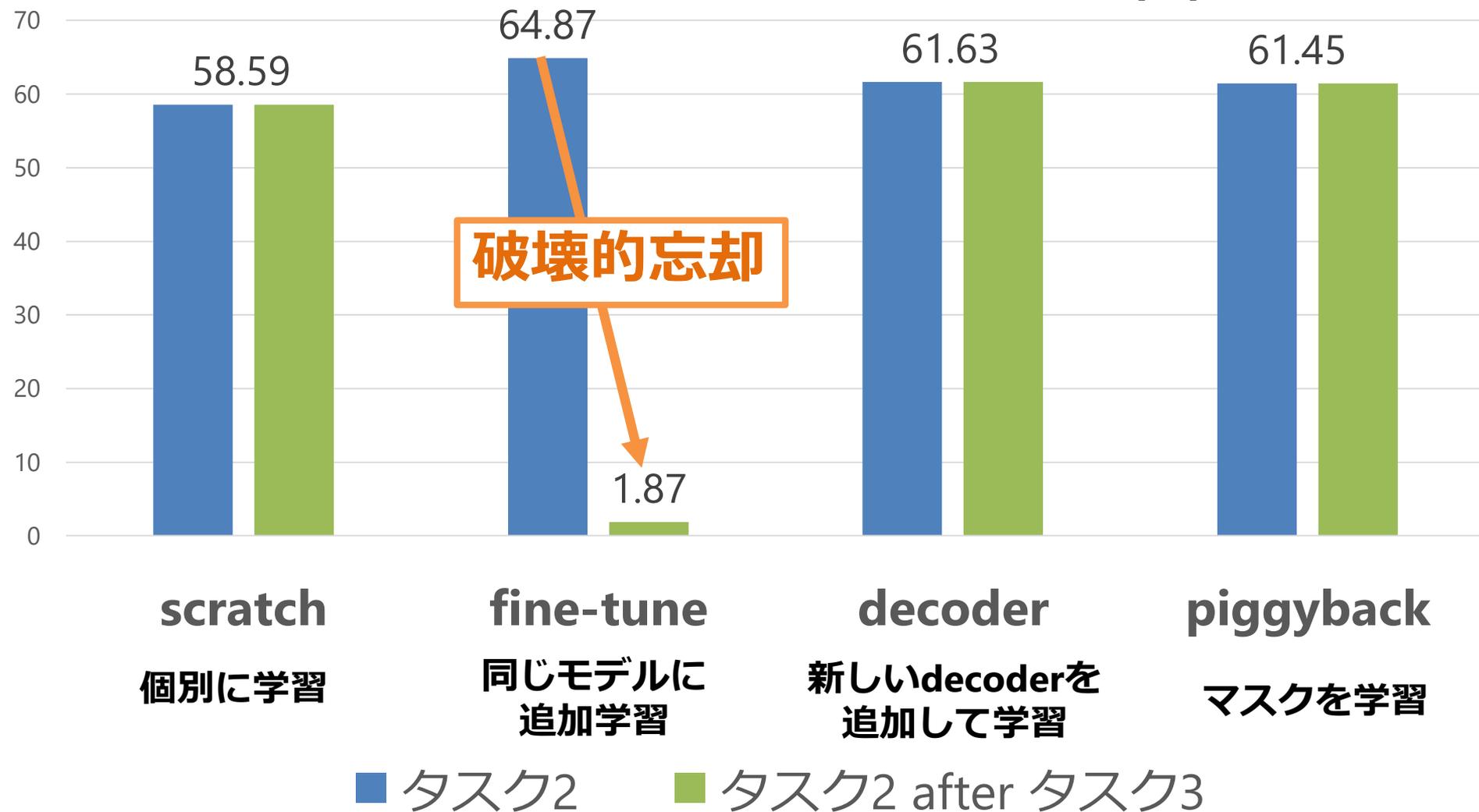
実験結果

タスク1 MS COCO 領域分割 mIoU (%)



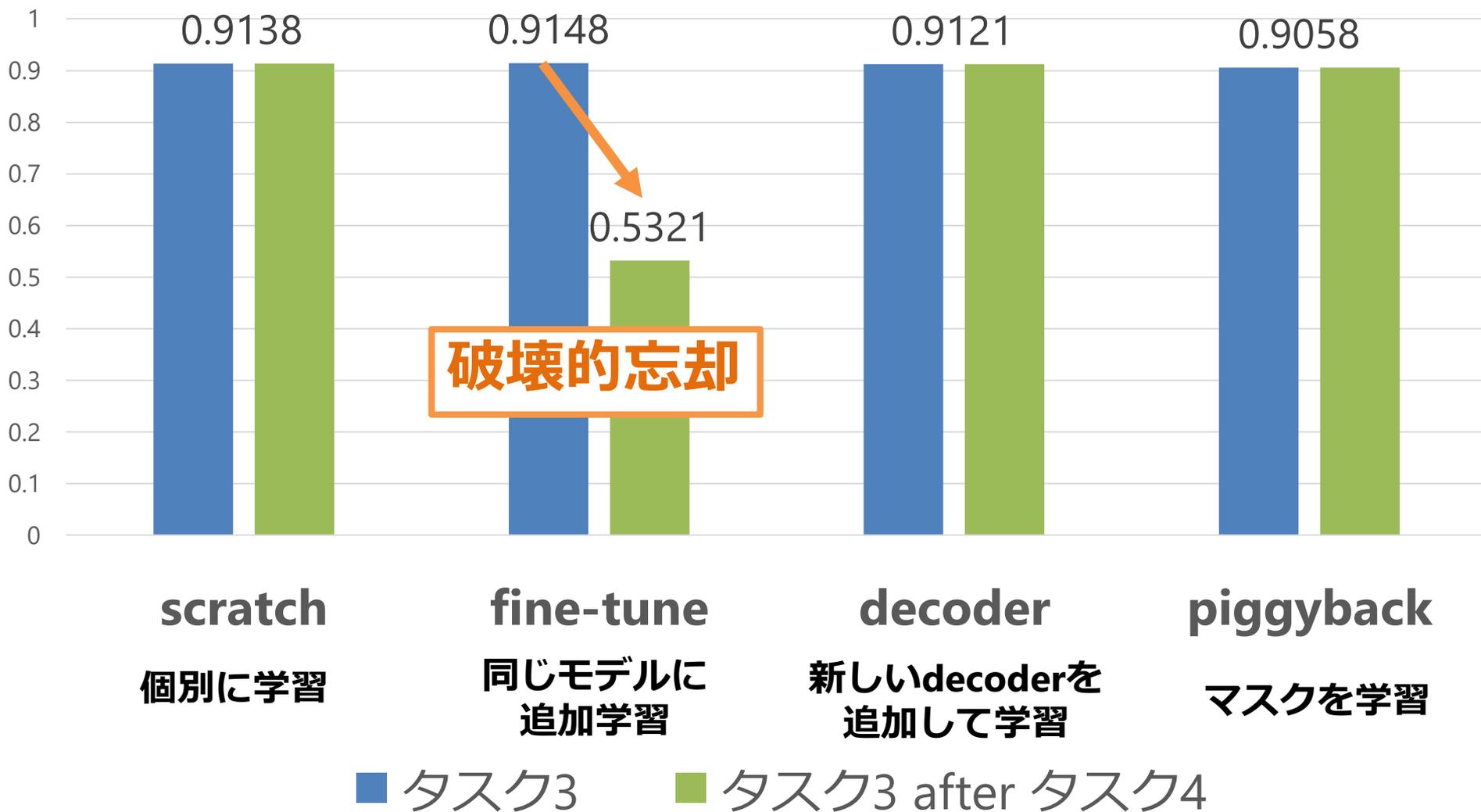
実験結果

タスク2 Pascal VOC 領域分割 mIoU (%)



実験結果

タスク3 濃淡画像着色 SSIM



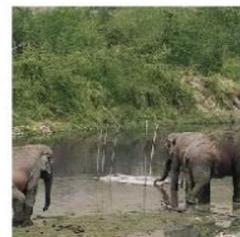
実験結果

● 生成例

タスク2



タスク3



タスク4



Input or Ground Truth

scratch

fine-tune

decoder

Piggyback

実験結果

● 各手法のモデルサイズ

	Scratch	Fine-tune	Decoder	Piggyback
モデルサイズ (MB)	282.0 (56.4×5)	56.4 (56.4×1)	138.4 (56.4+20.5*4)	63.6 (56.4+1.8*4)

破壊的忘却

● オーバヘッド

- CNN: 32 bit float
- バイナリーマスク: 1 bit

- 1タスクあたりのオーバヘッド: $\frac{1}{32} \times CNNsize$

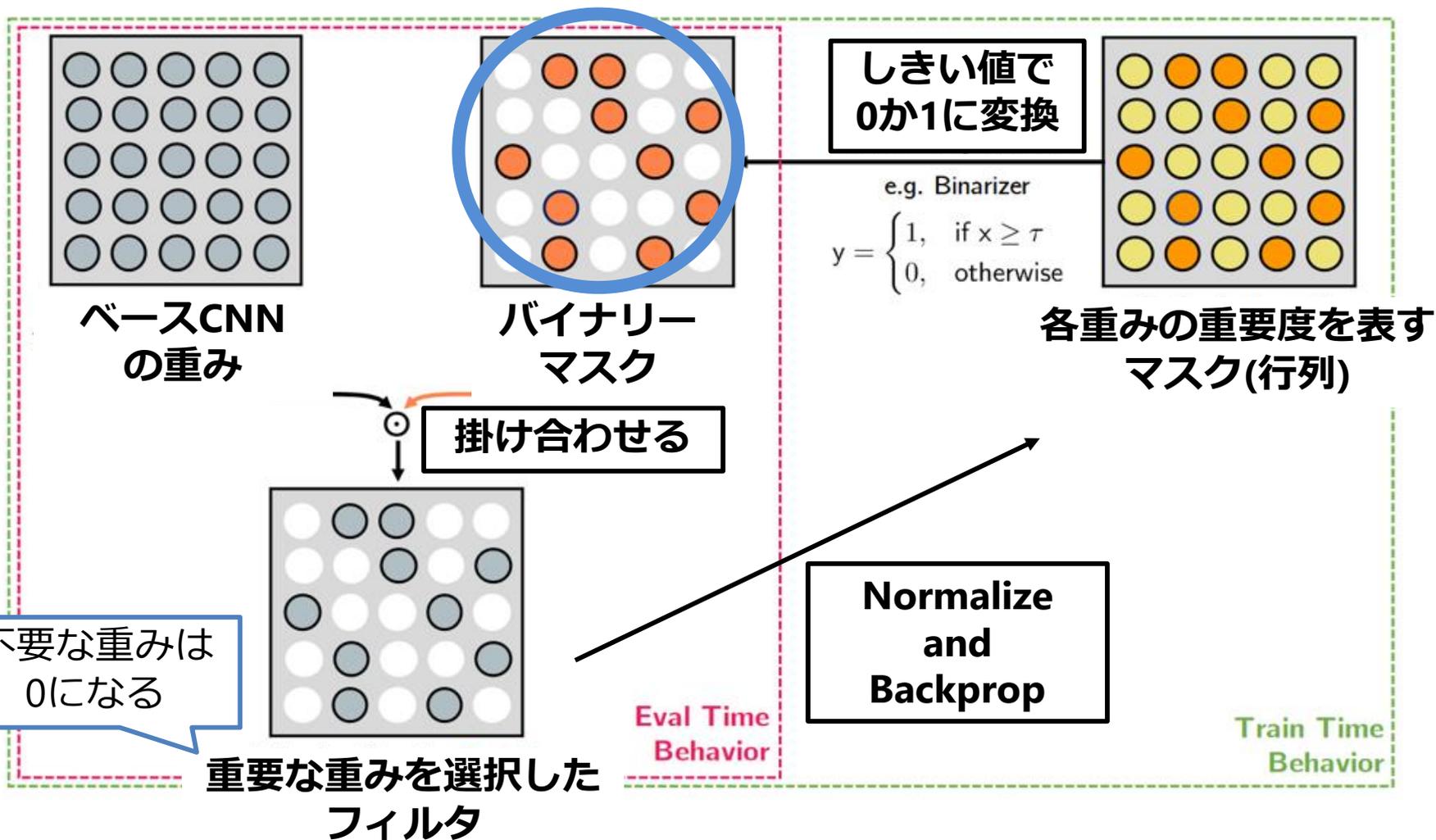
実験結果

- まとめ

Piggybackは

- ベースラインと同等の性能
- 破壊的忘却を回避
- 小さいモデルサイズ

[再掲] マスクの学習

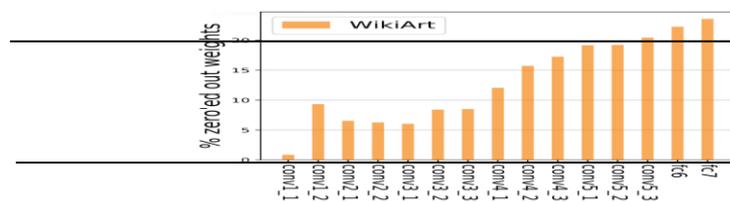


Piggyback バイナリーマスク

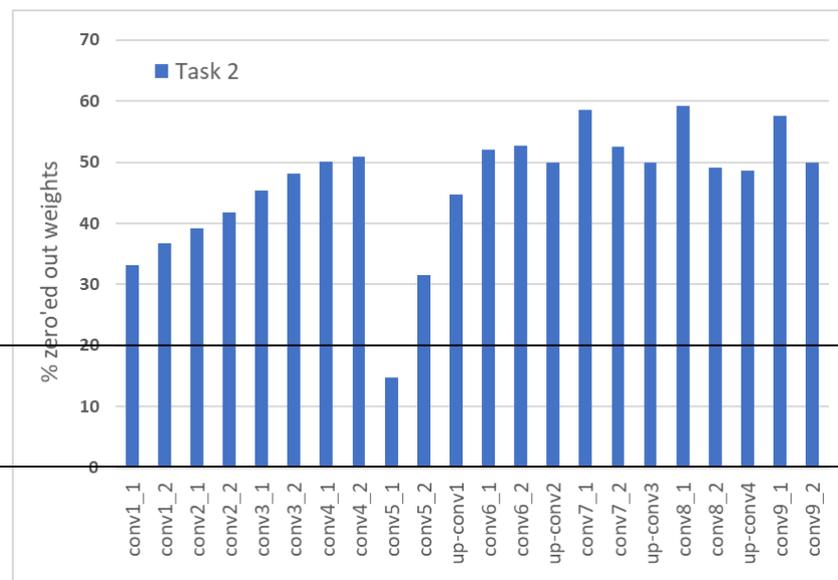
● クラス分類 と 画像変換(領域分割)

● 0の割合が全体的に大きい

- パラメーター数大
- アーキテクチャ



クラス分類



画像変換 (領域分割)

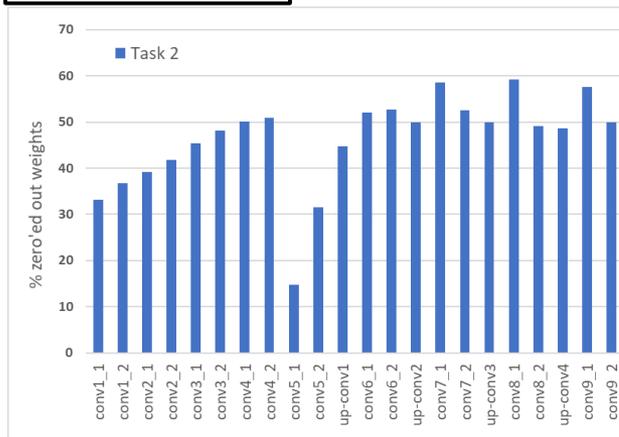
Piggyback バイナリーマスク

- 画像変換タスク
同士で比較

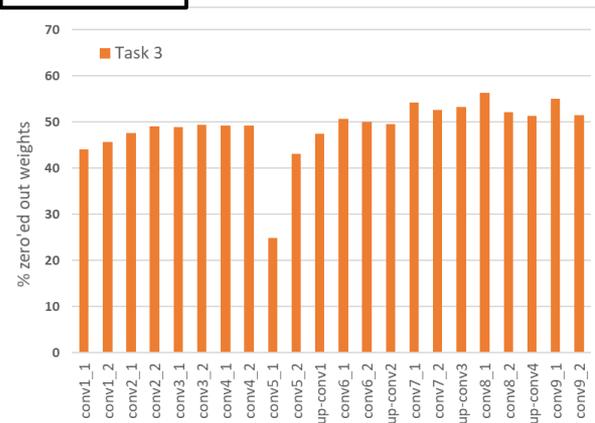
- 低レイヤー

- conv5_1

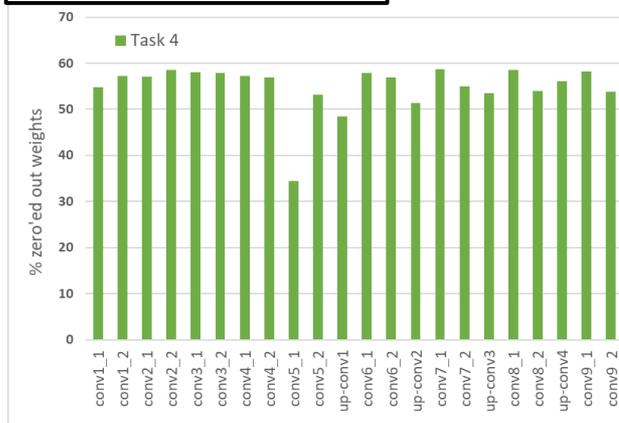
2 領域分割



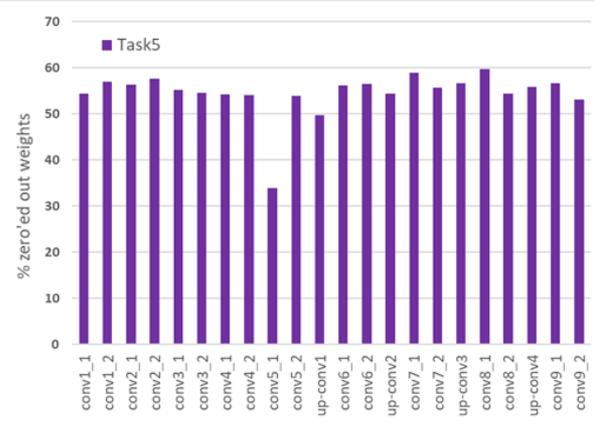
3 着色



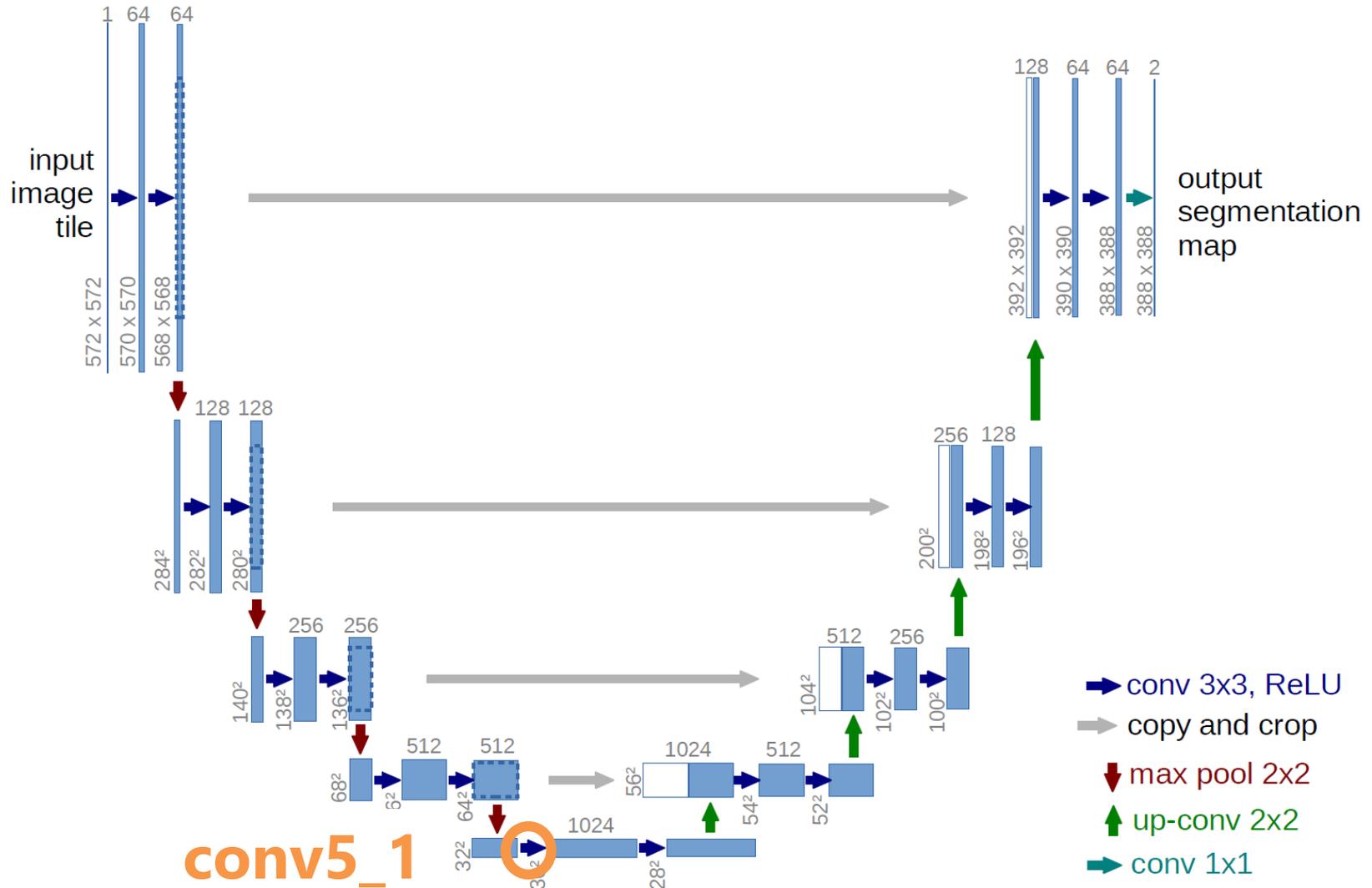
4 スタイル変換



5 スタイル変換



U Net



Piggyback バイナリーマスク

● バイナリーマスクの類似度

- 各タスクでの重要な重みは異なる

	タスク1 領域分割	タスク2 領域分割	タスク3 着色	タスク4 スタイル変換	タスク5 スタイル変換
タスク2	0.5075	-	-	-	-
タスク3	0.5042	0.5054	-	-	-
タスク4	0.4326	0.5034	0.5020	-	-
タスク5	0.4529	0.5029	0.5025	0.5210	-

まとめ

- 画像変換タスクでの連続学習を行った
- Piggyback はベースラインと同程度の性能を小さいオーバヘッドで達成

今後の課題

- 追加実験による汎用性の実証
- Piggyback の分析
- モデルの容量削減

WEC
TOKYO



Piggyback 有効性

なんで piggyback が効くの？

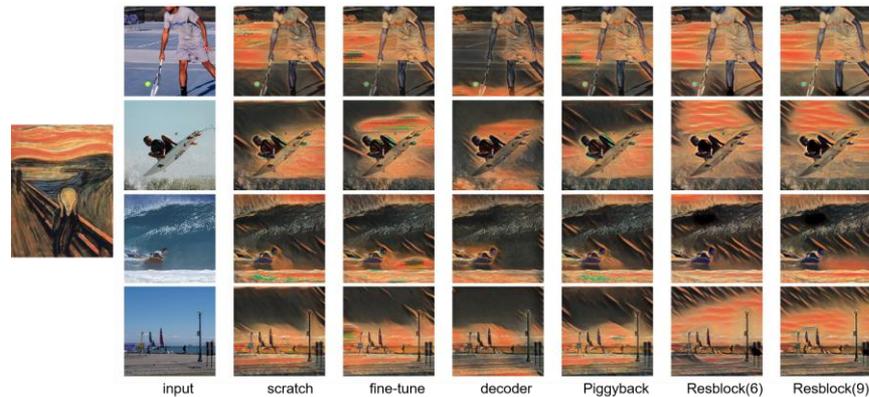
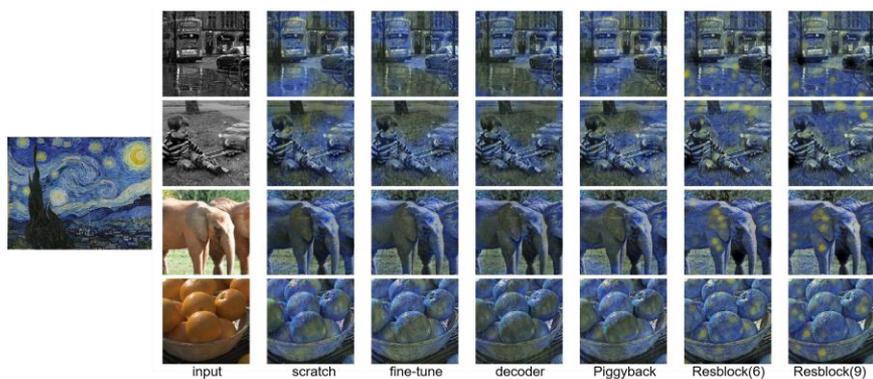
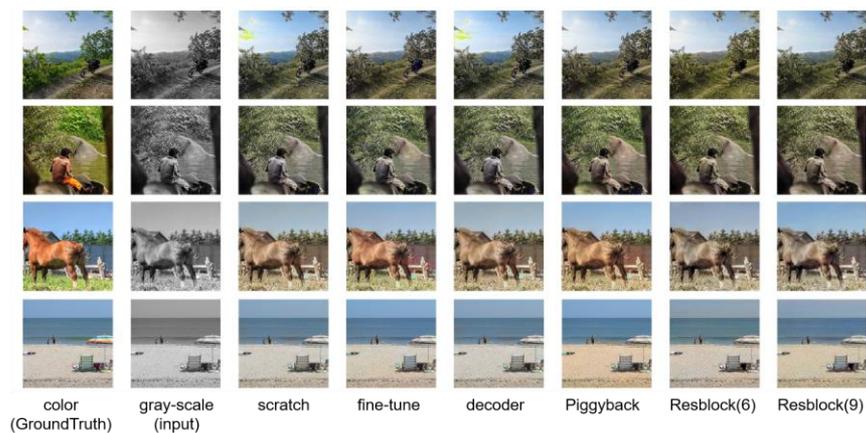
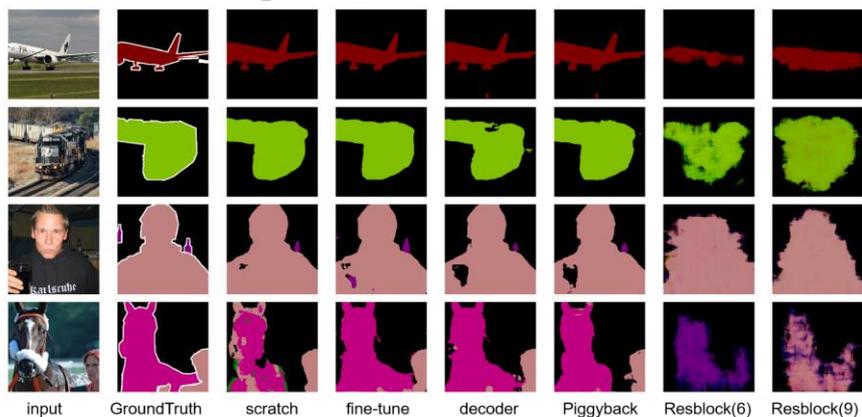
あくまでも予測ですか...

- 冗長性...? (CNNの不思議)
- 元から重要なパラメータは少ない
- バイナリーネットワーク
 - 最適化されたバイナリーが得られる
 - 大小が大事?
- BN が入ってるから絶対的な大きさは関係ない
 - 重みの組み合わせが重要 ?

	Scratch	Fine-tune	Decoder	Piggyback
タスク1 (mIoU(%))	21.47			
タスク2 (mIoU(%))	58.59	64.87	61.63	61.45
タスク3 (SSIM)	0.9138	0.9148	0.9121	0.9058
タスク4 (SSIM)	0.3678	0.3555	0.3595	0.3501
タスク5 (total loss)	447480	490490	544348	521476
タスク1 after タスク2	-	0.70	21.47	21.47
タスク2 after タスク3	-	1.87	61.63	61.45
タスク3 after タスク4	-	0.5321	0.9121	0.9058
モデルサイズ (MB)	282.0 (56.4×5)	282.0 (56.4)	138.4 (56.4+20.5*4)	63.6 (56.4+1.8*4)

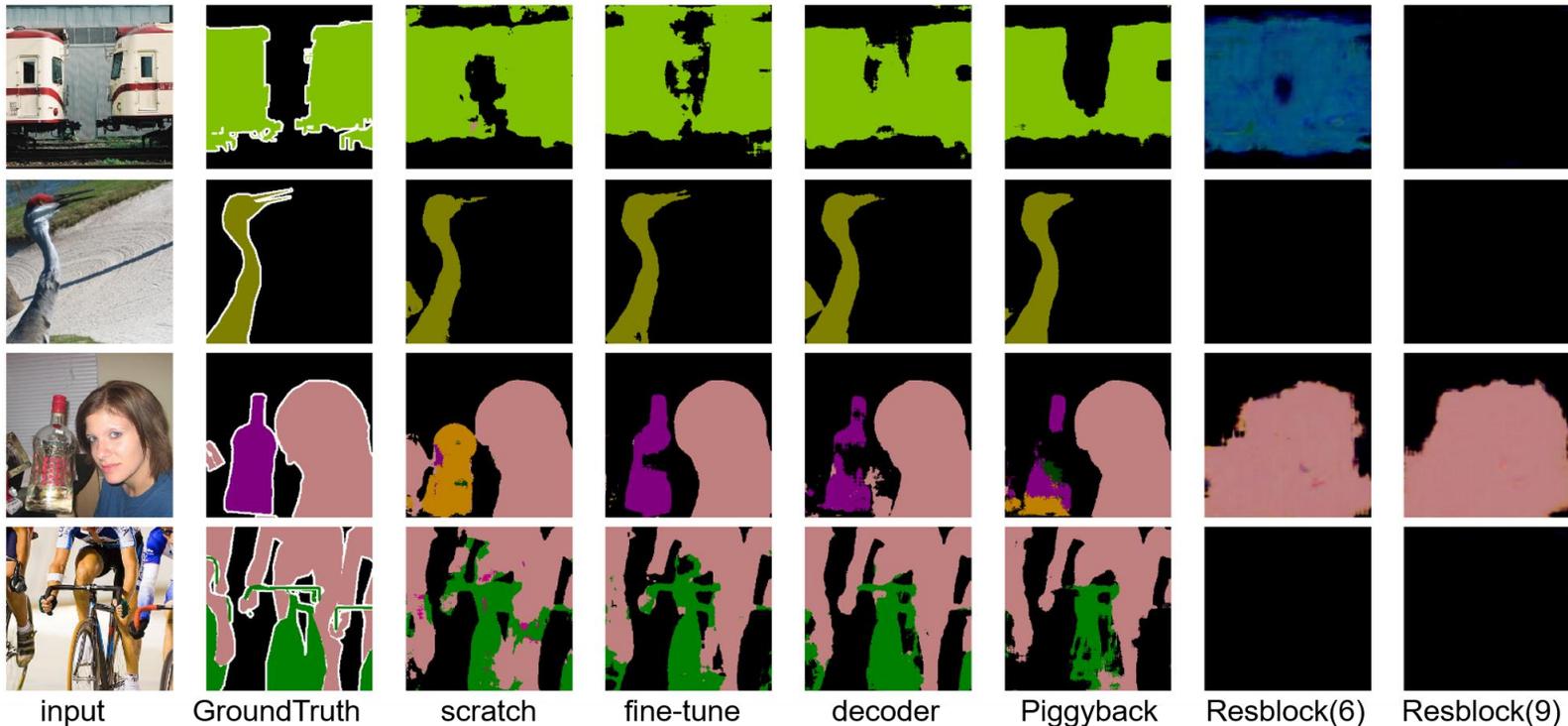
実験結果

● 生成結果



今後の課題

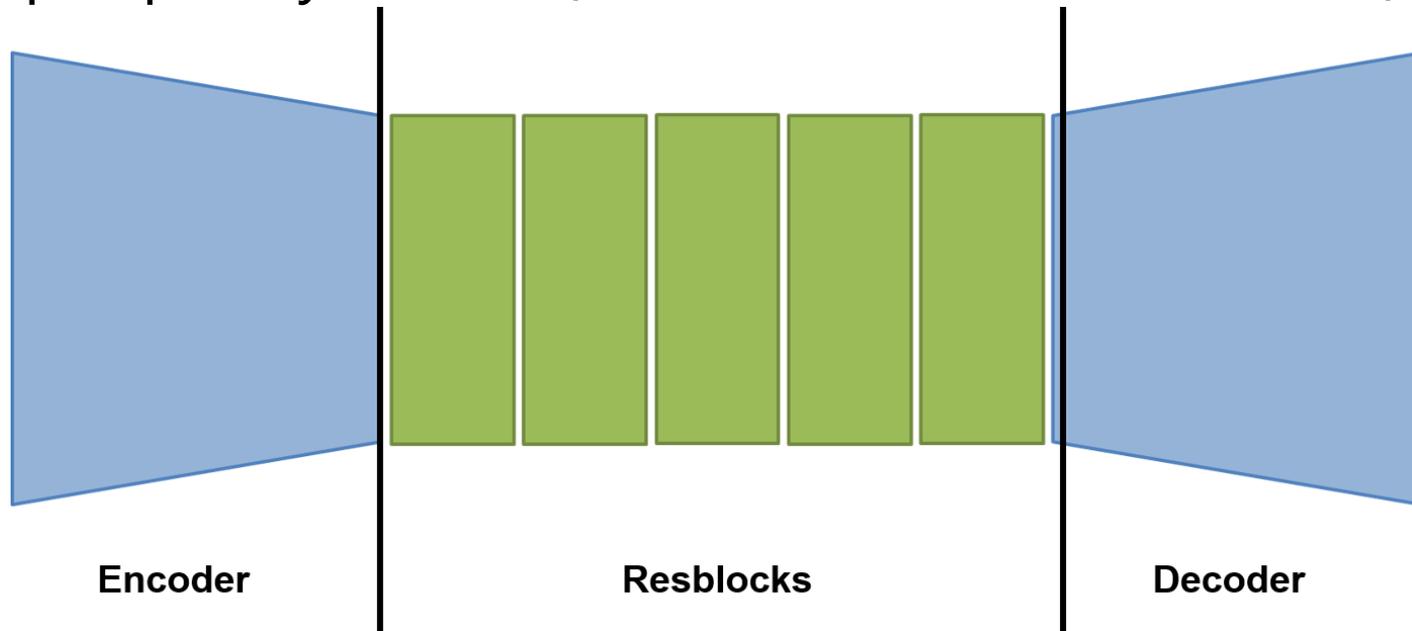
- Piggyback + Resblock
- 追加実験による汎用性の実証
- Resblock での領域分割



関連研究 : Residual Connection

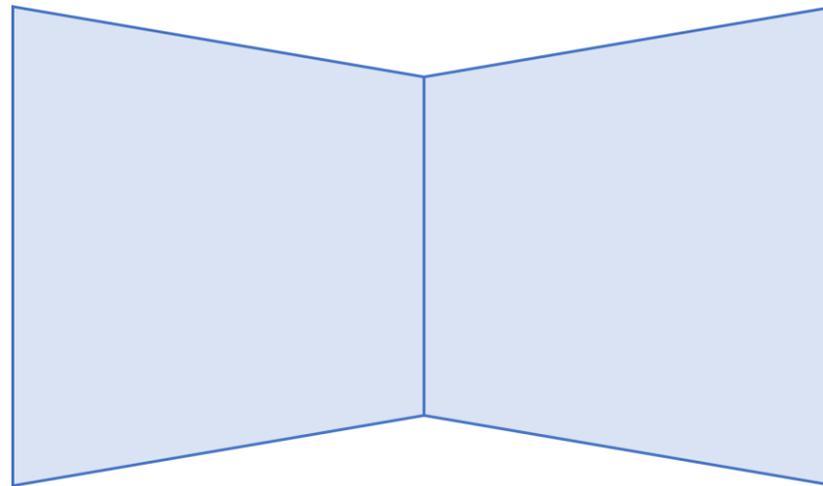
● Resblock付きEncoder-Decoder CNN

- Fast Neural Style Transfer (Johnsonら, ECCV2016)
- pix2pix, CycleGAN (Isola, Zhuら, CVPR2017, ICCV2017)



手法 : Resblock

- Auto Encoder + Resblock
- Auto Encoder : 入力と同じものを出力



特徴抽出

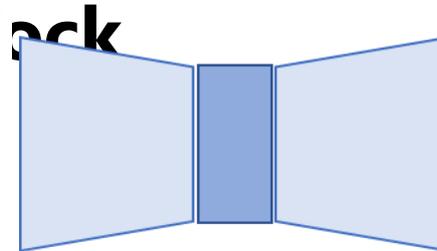
画像生成



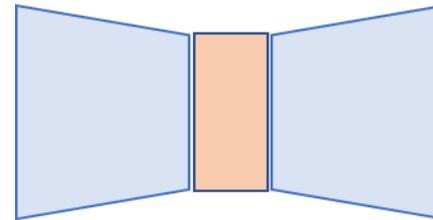
手法 : Resblock

- Auto Encoder

- タスク1



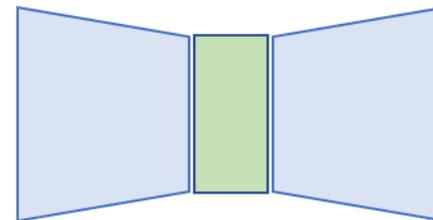
- タスク2



- タスク3



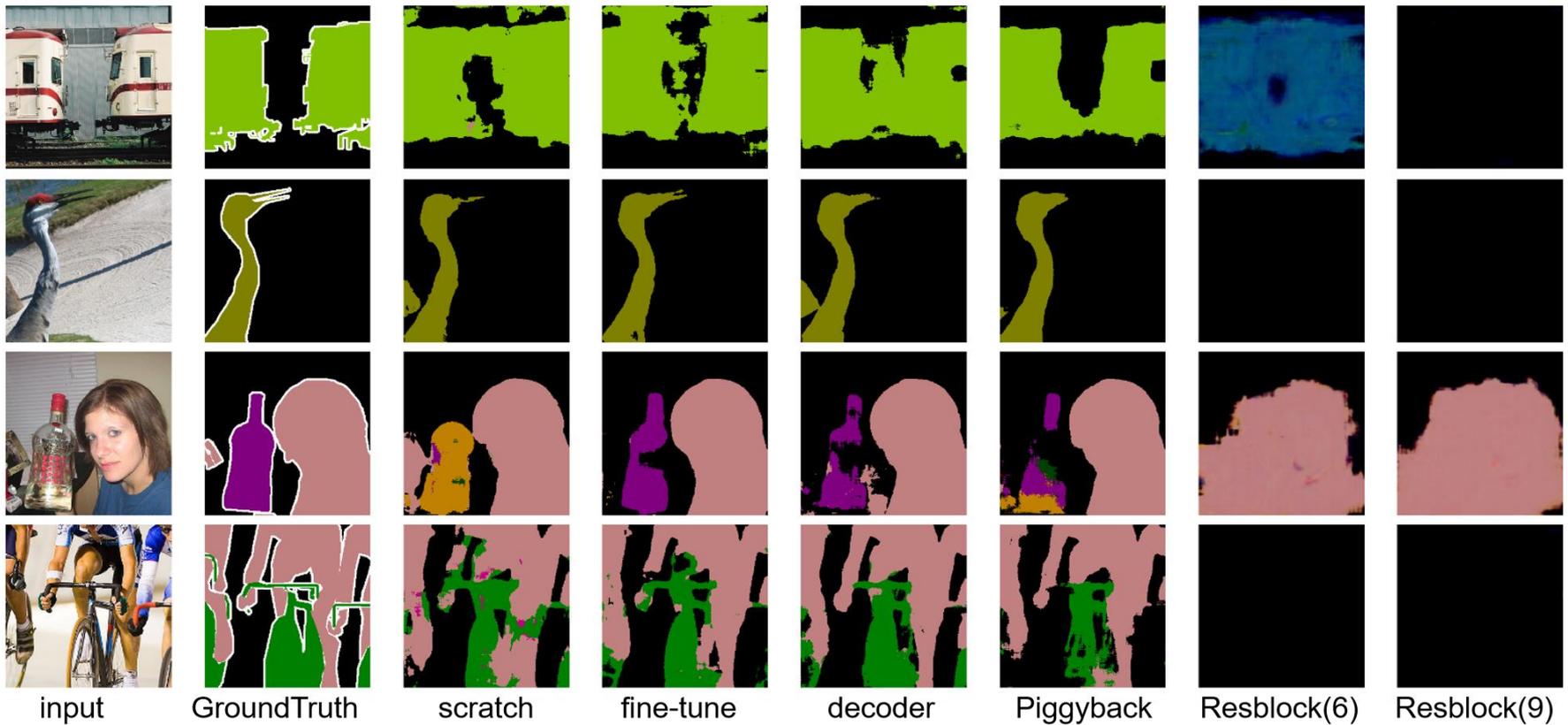
input



output

考察

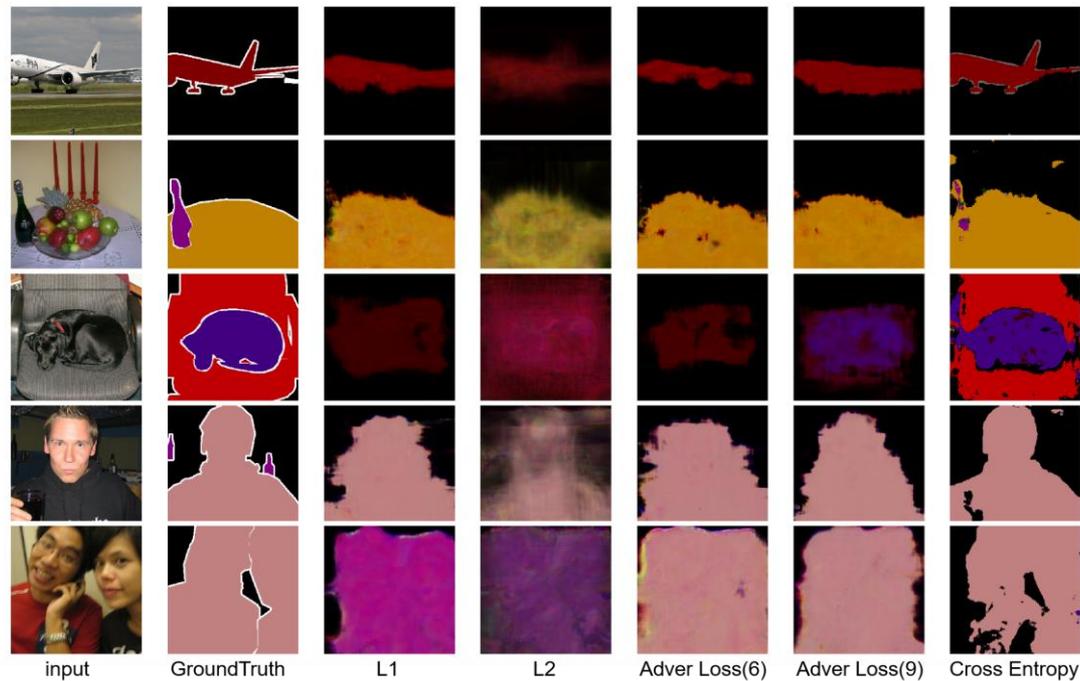
● Resblock 領域分割



考察

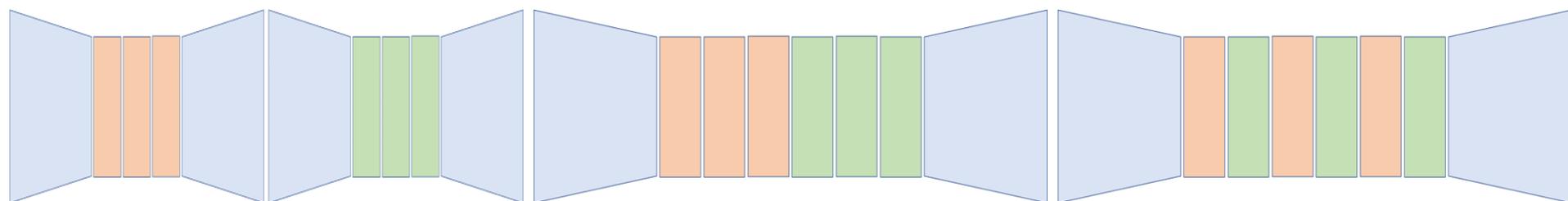
● Resblock 領域分割

	L1	L2	Adv Loss 6	Adv Loss 9	Cross Entropy
mIoU	7.11	4.49	4.26	9.54	52.55



考察

● Resblock 連結



Two Stage

Sequentially

Alternately



	Scratch	Fine-tune	Decoder	Piggyback	Resblock 6	Resblock 9
タスク1 (mIoU(%))	21.47				0.57	0.71
タスク2 (mIoU(%))	58.59	64.87	61.63	61.45	4.26	9.54
タスク3 (MSE, SSIM)	244.00 0.9138	237.92 0.9148	241.66 0.9121	242.49 0.9058	532.83 0.9281	527.13 0.9286
タスク4 (SSIM, total loss)	0.3678 413833	0.3555 405893	0.3595 473723	0.3501 528587	0.3467 460268	0.3524 442211
タスク5 (total loss)	447480	490490	544348	521476	520221	494600
タスク1 after タスク2	-	0.70	21.47	21.47	0.57	0.71
タスク2 after タスク3	-	1.87	61.63	61.45	4.26	9.54
タスク3 after タスク4	-	870.18 0.5321	241.66 0.9121	242.49 0.9058	532.83 0.9281	527.13 0.9286
モデルサイズ (MB)	282.0	282.0	138.4	63.6	683.7	1021.2

	Scratch	Fine-tune	Decoder	Piggyback	Resblock 6	Resblock 9
タスク1 (mIoU(%))	21.47				0.57	0.71
タスク2 (mIoU(%))	58.59	64.87	61.63	61.45	4.26	9.54
タスク3 (MSE, SSIM)	244.00 0.9138	237.92 0.9148	241.66 0.9121	242.49 0.9058	532.83 0.9281	527.13 0.9286
タスク4 (SSIM, total loss)	0.3678 413833	0.3555 405893	0.3595 473723	0.3501 528587	0.3467 460268	0.3524 442211
タスク5 (total loss)	447480	490490	544348	521476	520221	494600
タスク1 after タスク2	-	catastrophic forgetting 0.70	NOT catastrophic forgetting 21.47	21.47	0.57	0.71
タスク2 after タスク3	-	1.87	61.63	61.45	4.26	9.54
タスク3 after タスク4	-	870.18 0.5321	241.66 0.9121	242.49 0.9058	532.83 0.9281	527.13 0.9286
モデルサイズ (MB)	282.0	282.0	138.4	63.6	683.7	1021.2

	Scratch	Fine-tune	Decoder	Piggyback
タスク1 (mIoU(%))	21.47			
タスク2 (mIoU(%))	58.59	64.87	61.63	61.45
タスク3 (SSIM)	0.9138	0.9148	0.9121	0.9058
タスク4 (SSIM)	0.3678	0.3555	0.3595	0.3501
タスク5 (total loss)	447480	490490	544348	521476
タスク1 after タスク2	-	0.70	21.47	21.47
タスク2 after タスク3	-	1.87	61.63	61.45
タスク3 after タスク4	-	0.5321	0.9121	0.9058
モデルサイズ (MB)	282.0 (56.4×5)	282.0 (56.4)	138.4 (56.4+20.5*4)	63.6 (56.4+1.8*4)

catastrophic forgetting

NOT catastrophic forgetting

考察 : Piggyback バイナリーマスク

	タスク1 領域分割	タスク2 領域分割	タスク3 着色	タスク4 スタイル変換	タスク5 スタイル変換
タスク2	0.5075	-	-	-	-
タスク3	0.5042	0.5054	-	-	-
タスク4	0.4326	0.5034	0.5020	-	-
タスク5	0.4529	0.5029	0.5025	0.5210	-