



「次世代人工知能・ロボット中核技術開発」 (人工知能分野)中間成果発表会

- 人間と相互理解できる人工知能に向けて-

深層表現学習技術の研究開発

平成29年3月29日

国立大学法人 東京大学 国立研究開発法人 産業技術総合研究所 松尾 豊

国立研究開発法人 産業技術総合研究所 国立研究開発法人 新エネルギー・産業技術総合開発機構



背景



- 深層学習(Deep Learning)は、画像認識を代表する多くの認識系のタスクで従来法を上回る性能を示しているが、認識系以外のタスクにおいては、その性能はまだまだ高いわけではない。
- 実世界データへの応用においては、認識のみならず、行動・介 入の計画や制御が行えることが重要
- そこで、深層学習の技術を認識以外のタスクも扱えるように拡 張するための研究開発を行う。

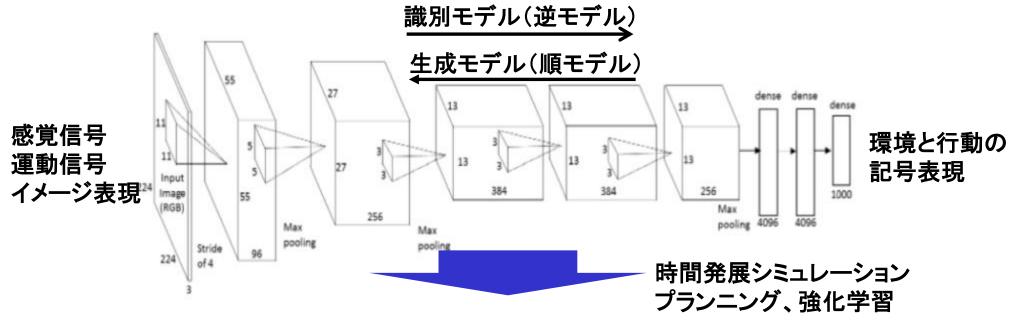


研究開発の内容



- 生成モデル系の深層学習技術と強化学習技術を組み合わせて、 記号とイメージの融合によるプラニングを可能にする新規な技 術を研究開発する。
- オブジェクトの認識だけでなく、認識したオブジェクト間の関係の

表現も獲得できるような深層学習の技術を研究開発する。





目標



最新の生成モデルによるauto-encoderの実装と、強化学習との組み合わせに関してのアルゴリズムの全体像が固まっていることを目標にする。また、簡単なロボットの試作を行い、さまざまな産業分野(例えば農業、建設、食品加工等)に活用することを目指したプロトタイプを1種類以上構築する。

先進中核モジュールとして、1種類以上の産業分野(例えば農業、 建設、食品加工等)の実フィールドでの実証評価をめざす。



これまでの成果



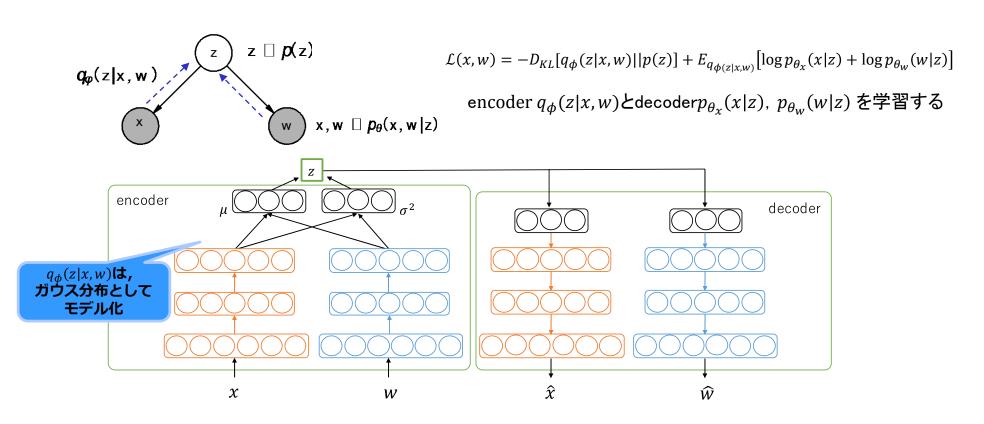
- 生成モデルの研究(マルチモーダルVAE): これを説明
 - ICLRに投稿
- 生成モデルによる翻訳の研究
 - ドイツ語⇔英語。Flickrデータセット等を用いる。
 - 画像を介した翻訳。いったん画像を生成してから画像とソース文で条件付けて、ターゲットの文を生成。少しだけ精度が上がる。
 - ICLRに投稿
- 密度比推定を用いた生成敵対的ネットワーク(GAN):
 - ICLRに投稿。arXivにアップすると、DeepMindから翌日引用される。
- SDAの高速化に関する研究: これを説明
 - 高速化手法は、深層学習研究の全般にわたって重要
- 深層強化学習の実世界への適用:
 - 建設用の機械に対して適用(詳細は企業側が発表できる段階に)



マルチモーダルVAE



2つのモーダルx, wに対して同時確率 p(x,w)をモデル化する。 (一方から他方を生成するモデルはあるが、平等に扱うものはなかった。)



本手法の利点

単独のモーダル情報よりも潜在変数のよりよい特徴表現が獲得できる. 同時分布p(x,w)をモデル化しているため、p(x)やp(x|w)なども求めることが可能. 2つ以上のモーダル情報にも容易に拡張可能.



CelebA dataset:画像とタグを複数のモーダルと考える



$x \rightarrow x$ の検証

周辺尤度p(x)の比較

モデル	テスト尤度	
VAE-GAN	-375.21	
MVAE-GAN	-292.69	

• MVAE-GANによる再構成画像



元画像



再構成画像

$w \rightarrow x$ **の検証**

モデル	テスト尤度	
CVAE-GAN	-392.07	
MVAE-GAN	-258.13	

- $q_{\phi_w}(z|w)$ の平均 μ の利用
 - 様々な属性wの値を設定して生成

























モナリザを笑わせる

生成したい画像



• モナリザの顔をxとし、 $w \sim p(w|x)$ によって属性を推定する





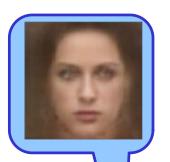
Male Eyeglasses Smiling · · ·
-0.29 -1 -0.98 · · ·

元画像x

推定した属性ル

• wとw* (Smiling=3にしたもの) で、平均的な顔を生成する

wの平均的な顔





w*の平均的な顔

それぞれのzに対してベクトルを計算し、オリジナルのものに加える

元画像

z空間



モナリザを笑わせる



• 属性「Smiling」を3にしたモナ・リザの顔



元画像

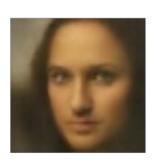


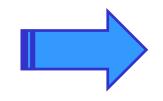
再構成画像



Smiling=3にした再構成画像

同様にして様々な属性を変化させることができる.





Male = 3







Eyeglasses = 1

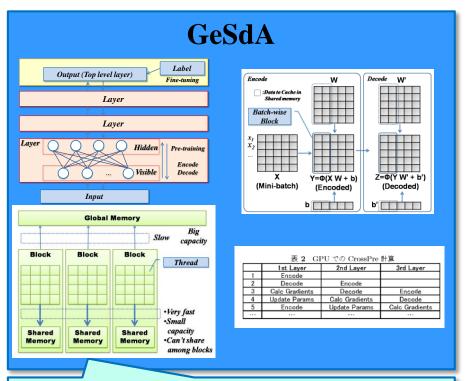


Mustache = 5



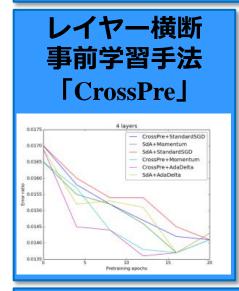
高速計算の枠組み





GPUでの並列化に適した Stacked denoising Autoencoderの開発

GPUによる高速Deep Learning 「GeSdA」の開発





情報処理学会論文誌 (Transaction of DB) 採録他

解決すべき技術課題

コンシューマレベルのGPU環境 へのDeep Learning技術の適用 GPU内での効率的な並列計算戦略 近似解を利用した

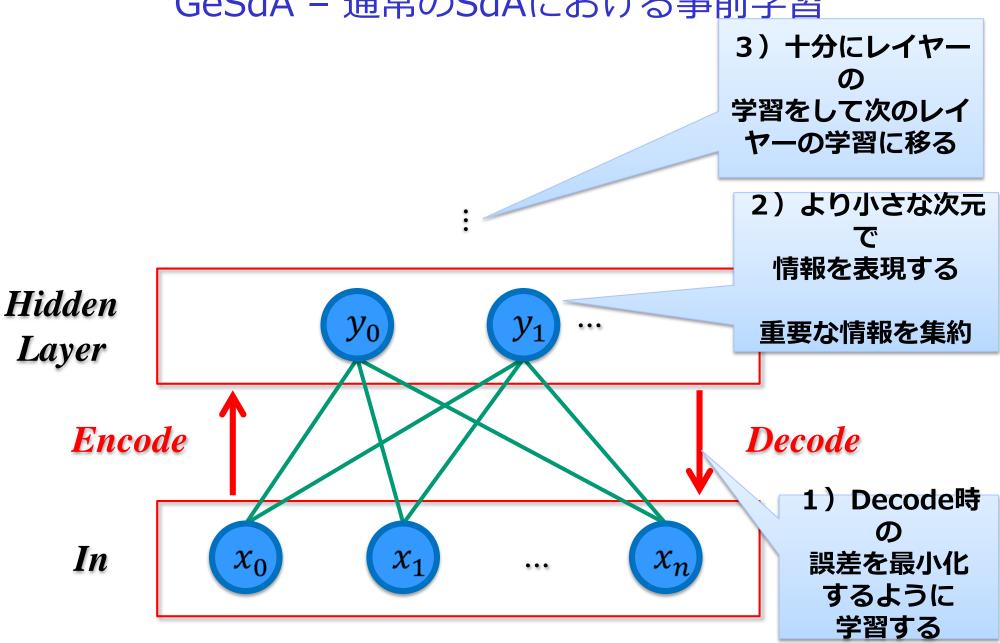
将来的なビジョン

コンシューマ環境や 組み込み機器などの 環境下でも高速に 動作可能なモデルの開発





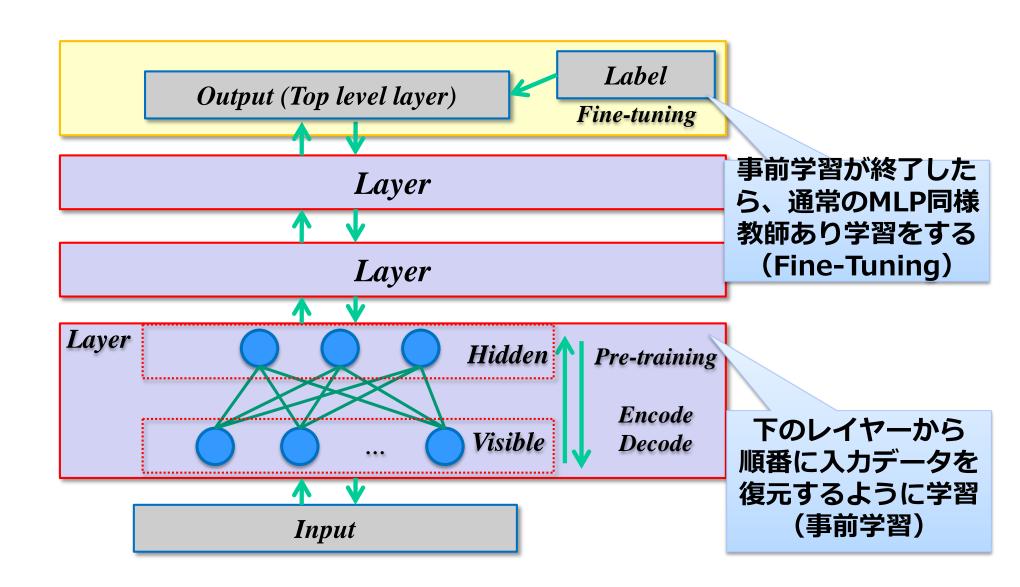
GeSdA - 通常のSdAにおける事前学習







GeSdA - 通常のSdAにおける事前学習







GeSdA - レイヤー横断事前学習手法「CrossPre」

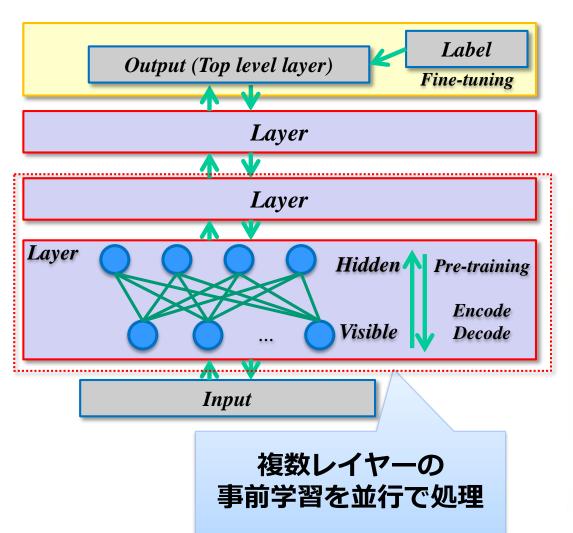
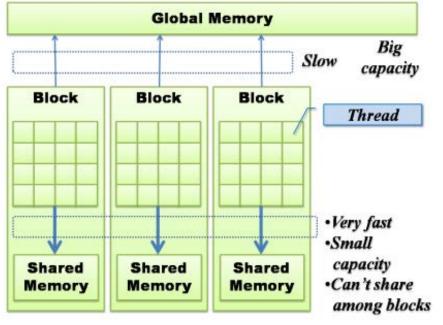


表 2 GPU での CrossPre 計算

	1st Layer	2nd Layer	3rd Layer
1	Encode		
2	Decode	Encode	
3	Calc Gradients	Decode	Encode
4	Update Params	Calc Gradients	Decode
5	Encode	Update Params	Calc Gradients
		•••	



DL+強化学習を産業機械へ活用

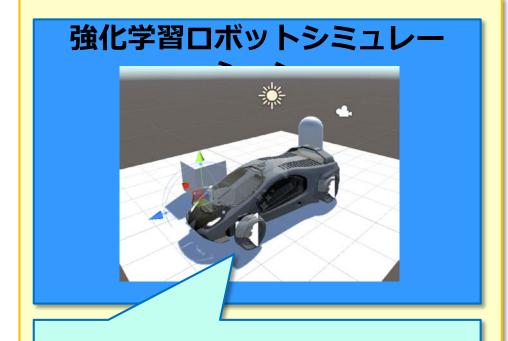


プロジェクトの概要

強化学習ロボットの開発



組み込み機器の制御に適した Deep Learning技術の研究開発 Deep Learningによるセンサ処理の高度化



ロボット制御のためのシミュレーション 環境の構築 無制限に学習可能な環境の構築

建設や農業分野における実例づくり





今後の内容

- 生成モデル、認識モデルと強化学習の組み合わせ方の検討
- 強化学習実験用ロボット構築
 - より高度なロボットプロトタイプの開発
- 深層強化学習の実世界への適用:
 - 建設用の機械等に対しての実装を進める