

AI 特許紹介(2)

～DeepMind 深層生成モデルにおける One-Shot Generalization 特許～

2019年5月10日

河野特許事務所

所長 弁理士 河野英仁

「AI 特許紹介」シリーズは、注目すべき AI 特許のポイントを紹介します。熾烈な競争となっている第4次産業革命下では AI 技術がキーとなり、この AI 技術・ソリューションを特許として適切に権利化しておくことが重要であることは言うまでもありません。

AI 技術は Google, Microsoft, Amazon を始めとした IT プラットフォーマ、研究機関及び大学から毎週のように新たな手法が提案されており、また AI 技術を活用した新たなソリューションも次々とリリースされています。

本稿では米国先進 IT 企業を中心に、これらの企業から出願された AI 特許に記載された AI テクノロジー・ソリューションのポイントをわかりやすく解説致します。

1.概要

特許権者 DeepMind Technologies Ltd

出願日 2017年2月3日

登録日 2019年1月8日

登録番号 US10176424

発明の名称 生成型ニューラルネットワーク

深層生成モデルを用いた One-Shot Generalization に関する発明である。

2.特許内容の説明

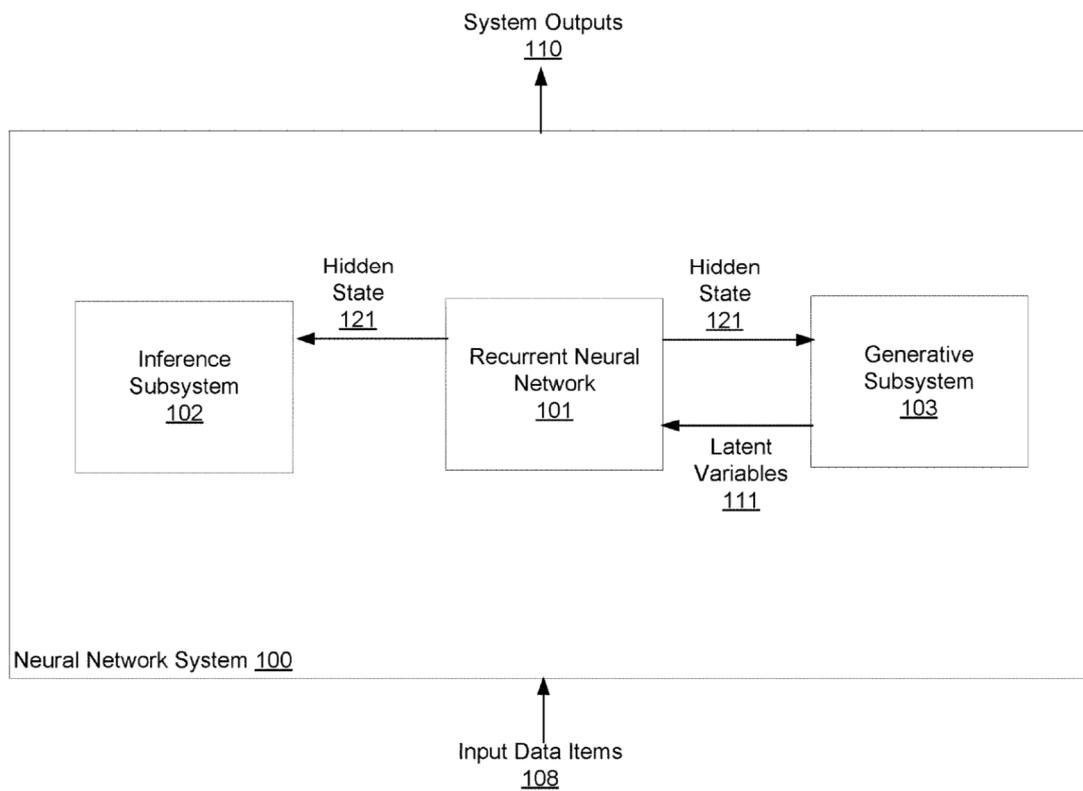
ニューラルネットワークシステム 100 は画像データ 108 の入力を受け付ける。ニューラルネットワークシステム 100 はリカレントニューラルネットワーク 101 及び生成サブシステム 103 を備え、生成サブシステムにより生成された画像をシステム出力 100 として出力する。

リカレントニューラルネットワーク 101 は、ネットワーク中間層の隠れ状態 121 を生成サブシステム 103 に出力する。生成サブシステム 103 は、所定数の時間ステップの各時間ステップについて、潜在変数 111 のセットを生成し、潜在変数のセットをリカ

レントニューラルネットワーク 101 へ出力する。

リカレントニューラルネットワーク 101 は生成サブシステム 103 から受け取った潜在変数を使用してリカレントニューラルネットワーク 101 の隠れ状態を更新する。生成サブシステム 103 は再び更新後の隠れ状態を使用して隠れキャンバスを更新する。

生成サブシステム 103 は、最後の時間ステップにおいて、最後の時間ステップで更新された隠れキャンバスを使用して出力イメージ 110 を生成する。



3.特許クレーム

424 特許クレーム 1 は以下の通りである。

1. 1つまたは複数のコンピュータによって実行されるニューラルネットワークシステムであって、

予め定められた数の時間ステップの各時間ステップに対して、その時間ステップに対する1組の潜在変数を受け取り、その1組の潜在変数を使用してリカレントニューラルネットワークの隠れ状態を更新するように構成されたリカレントニューラルネットワ

ークと、

生成サブシステムとを備え、

該生成サブシステムは、

所定数の時間ステップの各時間ステップについて：

該時間ステップのための潜在変数のセットを生成し、潜在変数のセットをリカレントニューラルネットワークへの入力として提供し、

リカレントニューラルネットワークの更新された隠れ状態を使用して隠れキャンバスを更新し、

所定の時間ステップ数の最後の時間ステップについては、

最後の時間ステップで更新された隠れキャンバスを使用して出力イメージを生成する。

4. One-Shot Generalization に関する研究

人間はあらゆるものを認識し、パーツに分け、そして新たなモノ、アイデアを創造する。人間は文字を創造し、また馬車から自動車、鳥から飛行機を創造するに至った。

この人間の脳の創造メカニズムを追求した人工知能が **One-Shot Generalization** である。DeepMind 社は **One-Shot Generalization** に深層学習を組み込む本案を提案した¹。

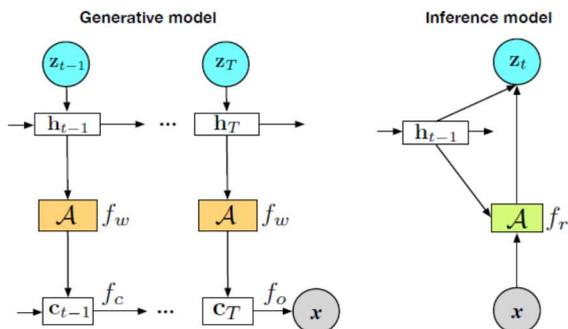
$$\text{Latent variables } \mathbf{z}_t \sim \mathcal{N}(\mathbf{z}_t | \mathbf{0}, \mathbf{I}) \quad t = 1, \dots, T \quad (3)$$

$$\text{Context } \mathbf{v}_t = f_v(\mathbf{h}_{t-1}, \mathbf{x}'; \theta_v) \quad (4)$$

$$\text{Hidden state } \mathbf{h}_t = f_h(\mathbf{h}_{t-1}, \mathbf{z}_t, \mathbf{v}_t; \theta_h) \quad (5)$$

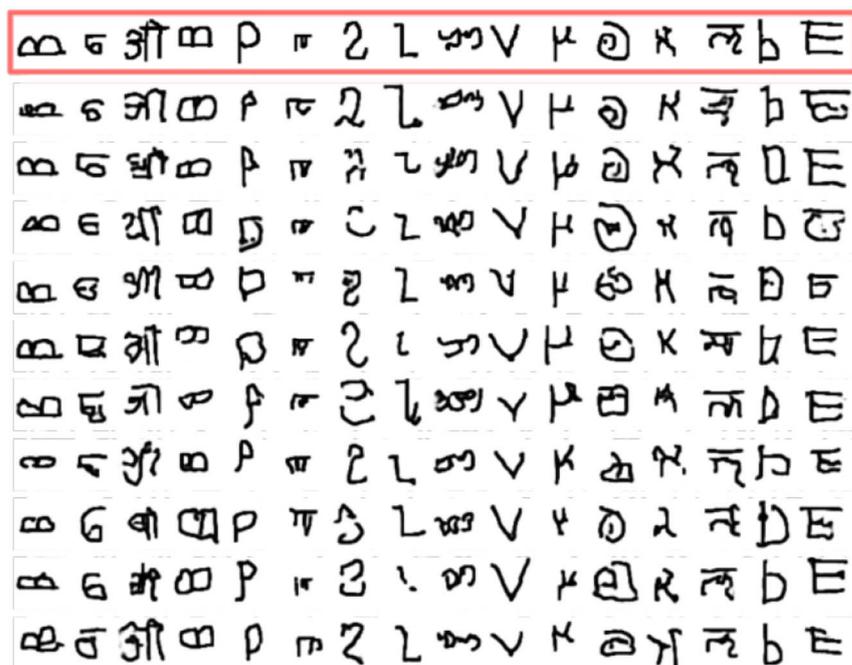
$$\text{Hidden Canvas } \mathbf{c}_t = f_c(\mathbf{c}_{t-1}, \mathbf{h}_t; \theta_c) \quad (6)$$

$$\text{Observation } \mathbf{x} \sim p(\mathbf{x} | f_o(\mathbf{c}_T; \theta_o)) \quad (7)$$



¹ Danilo J. Rezende “One-Shot Generalization in Deep Generative Models” 2016年5月26日

生成モデルは一般に上記図で表現でき、また上記式(3)~(7)により表現することができる。リカレントニューラルネットワークの中間層の隠れ状態 h_t は式(5)により表すことができ、モデルの隠れキャンバス C_t は式(6)で表すことができる。潜在変数の生成と隠れキャンバスの更新が繰り返し行われる。この繰り返しは人間の「思考時間」に該当する。そして、最終的に(7)で示す画像 x が出力される。



上記図の最初の行はテスト画像を示し、次の 10 行はモデルからの One-Shot サンプルを示す。



上記図²は空間変換及び追加キャンバス(32Steps)のモデルを使用して Multi-PIE デ

² YouTube より 2019 年 5 月 3 日 https://www.youtube.com/watch?v=6S6Tx_OtvnA

ータセット用に生成されたサンプルを示す。Multi-PIE データセットは、さまざまな視点からの 48×48 RGB 顔画像を含んでいる。いずれも新たな文字、顔画像が生成されている。

著者紹介

河野英仁

河野特許事務所、所長弁理士。立命館大学情報システム学博士前期課程修了、米国フランクリンピアースローセンター知的財産権法修士修了、中国清華大学法学院知的財産夏季セミナー修了、MIT(マサチューセッツ工科大学)コンピュータ科学・AI 研究所 AI コース修了。

[AI 特許コンサルティング](#)の他、米国・中国特許の権利化・侵害訴訟を専門としている。著書に「世界のソフトウェア特許(共著)」、「FinTech 特許入門」、「[AI/IoT 特許入門](#)」がある。

以上