

AI 特許紹介(84)
AI 特許を学ぶ！究める！
～Perfusion 特許～

2026 年 1 月 9 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

特許出願人 Nvidia

出願日 2023 年 10 月 31 日

公開日 2024 年 7 月 25 日

公開番号 US20240249446

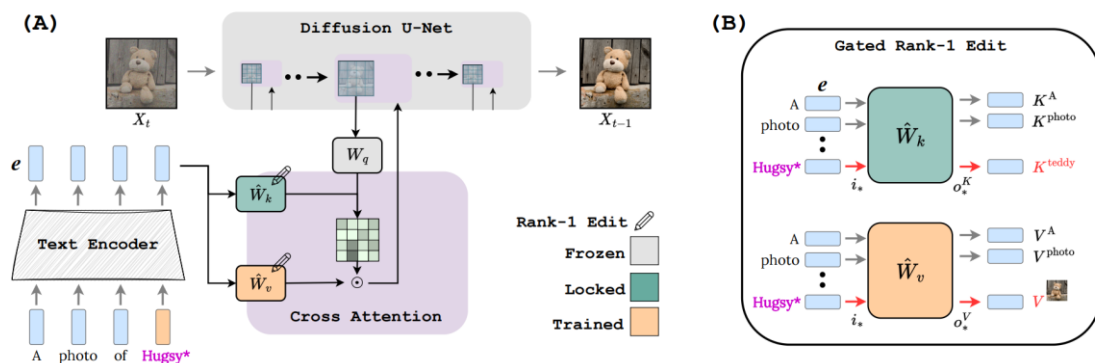
発明の名称 コンポーネントロックとランク 1 編集を備えたテキストから画像への拡散モデル

446 特許は、拡散モデルを用いたテキストから画像への変換において、新しい概念のクロスアテンションキーを上位カテゴリに「ロック」することで、過学習を回避するとともに、事前学習済みのテキスト画像拡散モデルの特定の層の重みを更新するランク 1 編集を用いることで、モデルをユーザによって入力された新しい概念に合わせてパーソナライズすることが可能な Perfusion 技術に関する。

2.特許内容の説明

下記図はテキストから画像への拡散モデルをパーソナライズするためのアーキテク

チャのブロック図を示す。



上記図(A)に示すように、プロンプトはテキストエンコーダにより、一連のエンコーディング e に変換される。各エンコーディング e は、拡散 U-Net デノイザーの一連のクロスアテンションモジュール（紫色のブロック）に送られる。拡大表示された紫色のモジュールは、キーとバリューの経路がテキストエンコーディングによってどのように条件付けられるかを示している。キーはアテンションマップを駆動し、アテンションマップはバリュー経路を調整する。

右側の図(B)はゲート付きランク 1 編集を示すブロック図である。上側のブロックにおいて、 K 経路はロックされているため、 \hat{W}_k に到達する e_{Hugsy} のエンコーディングはすべて、上位カテゴリ K^{teddy} のキーにマッピングされる。下側のブロックにおいて、 \hat{W}_v に到達する e_{Hugsy} のエンコーディングはすべて、学習された V^{Hugsy} にマッピングされる。

3.クレーム

446 特許のクレーム 1 は以下の通りである。

1. 以下のステップを含む方法：

デバイスにおいて：

事前学習済みのテキストから画像への拡散モデルによって学習される概念の少なくとも 1 つの画像を入力として受信し、前記概念は、前記テキストから画像への拡散モデルが事前学習された概念の上位カテゴリに関連付けられており、

前記概念の新しい画像を生成するために、前記テキストから画像への拡散モデルを学習し、以下のステップを含み、

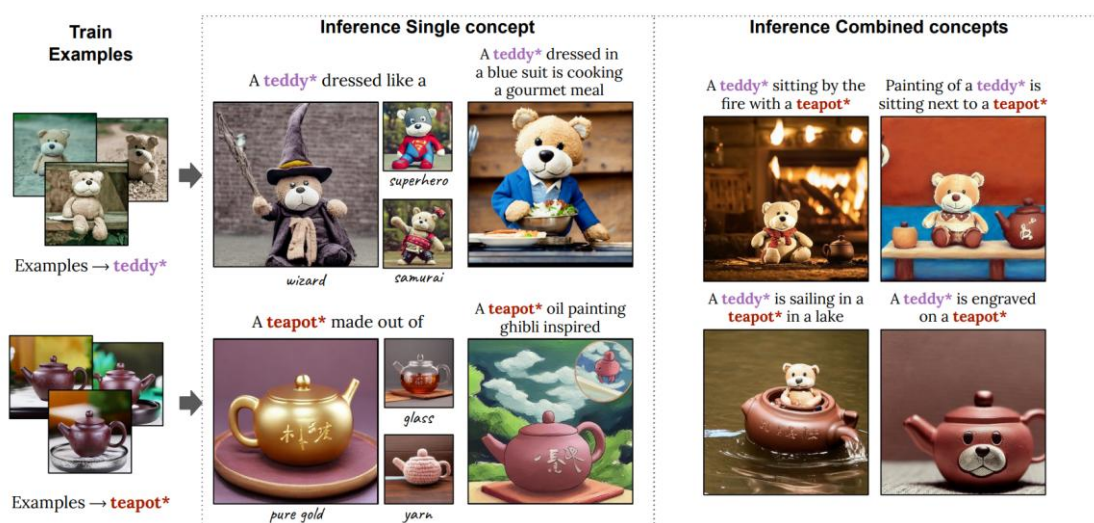
前記上位カテゴリに対する前記テキストから画像への拡散モデルの少なくとも 1 つの選択コンポーネントの活性化を計算し、

前記少なくとも 1 つの選択コンポーネントの活性化をロックしながら、前記概念の前記少なくとも 1 つの画像に基づいて、前記テキストから画像への拡散モデルの重みを

更新する。

4. 本特許に関連する論文

本特許に関する論文 “Key-Locked Rank One Editing for Text-to-Image Personalization”¹が、NVIDIA の YOAD TEWEL 氏らにより公表されている。



上記図は、パーソナライズされたテキストから画像への拡散モデルの例示的な入力と出力を示す。これらの画像は、パーソナライズされた概念のアイデンティティを損なうことなく、外観、ポーズ、およびコンテキストを大きく変更して生成できる。パーソナライズされたテキストから画像への拡散モデルは、コンパクト（例えば、概念ごとにわずか 100KB）である。

右側の図に示すように、Perfusion は、学習した概念を推論時に組み合わせて、複数の概念を並べて描写するシーンを作成し、また概念間の相互作用を作成することができる。Perfusion のプロジェクトは下記 WEB サイトより閲覧することができる。

<https://research.nvidia.com/labs/par/Perfusion/>

以上

著者紹介
河野英仁

¹ YOAD TEWEL, et al. “Key-Locked Rank One Editing for Text-to-Image Personalization” arXiv:2305.01644v2 [cs.CV] 5 Jun 2024

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、生成 AI ビジネスコース修了。

[AI 特許コンサルティング](#)、[医療 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 3](#)」、「[ブロックチェーン 3.0](#)(共著)」がある。