

【公報種別】特許法第 17 条の 2 の規定による補正の掲載
【部門区分】第 6 部門第 3 区分
【発行日】令和 4 年 5 月 31 日(2022.5.31)

【国際公開番号】WO2019/221985
【公表番号】特表 2021-524099(P2021-524099A)
【公表日】令和 3 年 9 月 9 日(2021.9.9)
【出願番号】特願 2020-564186(P2020-564186)
【国際特許分類】

G 0 6 N 20/00(2019.01)

10

G 0 6 T 7/00(2017.01)

【F I】

G 0 6 N 20/00

G 0 6 T 7/00 3 5 0 C

【手続補正書】

【提出日】令和 4 年 5 月 23 日(2022.5.23)

【手続補正 1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

20

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

第 1 モダリティからの入力データおよび前記第 1 モダリティとは異なる第 2 モダリティからの入力データを含む複数のモダリティからの入力データを受信するように構成されたマルチモーダル統計モデルを訓練する方法であって、前記方法は

前記第 1 モダリティの訓練データおよび前記第 2 モダリティの訓練データを含む訓練データにアクセスする、訓練データアクセス工程と、

前記マルチモーダル統計モデルを、前記訓練データを用いて訓練する、訓練工程であって、 30

前記マルチモーダル統計モデルは、前記第 1 モダリティおよび前記第 2 モダリティの入力データをそれぞれ処理する第 1 エンコーダおよび第 2 エンコーダと、第 1 モダリティ埋め込みおよび第 2 モダリティ埋め込みと、共同モダリティ表現と、予測子とを含む複数の構成要素を含み、前記第 1 モダリティ埋め込みおよび前記第 2 モダリティ埋め込みは、前記共同モダリティ表現を第 1 モダリティの空間および第 2 モダリティの空間にそれぞれ投影するための複数のベクトルおよび複数の重みを含み、前記訓練工程は、

前記訓練データを使用して、前記複数の構成要素のパラメータの値を推定することを含む訓練工程と、

前記マルチモーダル統計モデルの前記複数の構成要素のパラメータの推定した前記値を記憶することにより、前記マルチモーダル統計モデルを指定する情報を少なくとも部分的に記憶する工程と、を備える方法。 40

【請求項 2】

前記訓練工程は、前記第 1 モダリティ埋め込みおよび前記第 2 モダリティ埋め込みと前記共同モダリティ表現と予測子とのパラメータの値を推定する前に前記第 1 エンコーダおよび前記第 2 エンコーダのパラメータの値を推定する工程をさらに含む、請求項 1 に記載の方法。

【請求項 3】

前記訓練工程は、前記第 1 モダリティおよび前記第 2 モダリティの第 1 デコーダおよび第 2 デコーダのパラメータの値を推定する工程をさらに含む、請求項 2 に記載の方法。

【請求項 4】

50

前記訓練工程は、前記第 1 エンコーダおよび前記第 2 エンコーダのパラメータの値の推定を、前記共同モダリティ表現のパラメータの値の推定と共同とする工程をさらに含む、請求項 1 に記載の方法。

【請求項 5】

前記訓練工程は、前記第 1 モダリティの第 1 デコーダおよび前記第 2 モダリティの第 2 デコーダのパラメータの値を推定する工程をさらに含む、請求項 4 に記載の方法。

【請求項 6】

前記訓練工程は、

前記第 1 モダリティの前記訓練データにおける第 1 データ入力にアクセスする工程と、
前記第 1 データ入力を前記第 1 エンコーダに提供して、第 1 特徴ベクトルを生成する工程と、

10

前記共同モダリティ表現、前記第 1 モダリティ埋め込み、および前記第 1 特徴ベクトルを使用して、第 2 特徴ベクトルを特定する、第 2 特徴ベクトル特定工程と、

前記第 2 特徴ベクトルを入力として第 1 デコーダに提供して、第 1 データ出力を生成する工程と、を含む請求項 1 に記載の方法。

【請求項 7】

前記第 1 データ出力を前記第 1 データ入力と比較する工程と、

前記比較の結果に基づき、前記共同モダリティ表現の 1 つ以上のパラメータの 1 つ以上の値を更新する工程と、をさらに含む請求項 6 に記載の方法。

【請求項 8】

20

前記訓練工程は、

前記第 1 モダリティ埋め込みおよび前記第 2 モダリティ埋め込みと前記共同モダリティ表現とのパラメータの値を推定することにより、少なくとも部分的に第 1 訓練段階を実行する、第 1 訓練段階実行工程および、

前記予測子のパラメータの値を推定することにより、少なくとも部分的に第 2 訓練段階を実行する、第 2 訓練段階実行工程、を含む請求項 1 に記載の方法。

【請求項 9】

前記第 1 エンコーダは d 次元ベクトルを出力するように構成され、前記共同モダリティ表現は N 個の m 次元ベクトルを含み、前記第 1 モダリティ埋め込みは $m * d$ の重みを含む、請求項 6 に記載の方法。

30

【請求項 10】

前記第 2 特徴ベクトル特定工程は、

前記第 1 モダリティ埋め込みを使用することにより、前記共同モダリティ表現を前記第 1 モダリティの空間に投影して、 N 個の d 次元ベクトルを取得する工程と、

前記共同モダリティ表現における前記 N 個の d 次元ベクトルの中から、類似性メトリックに従って前記第 1 特徴ベクトルに最も類似する第 3 特徴ベクトルを特定する工程と、

前記第 1 特徴ベクトルを前記第 3 特徴ベクトルと集約することにより前記第 2 特徴ベクトルを生成する工程と、を含む請求項 9 に記載の方法。

【請求項 11】

前記第 2 特徴ベクトル特定工程は、

40

前記第 1 モダリティ埋め込みを使用することにより、前記共同モダリティ表現を前記第 1 モダリティの空間に投影して、 N 個の d 次元ベクトルを取得する工程と、

前記 N 個の d 次元ベクトルの少なくとも一部と前記第 1 特徴ベクトルとの間の類似性に従って、前記共同モダリティ表現における前記 N 個の d 次元ベクトルの前記少なくとも一部の重みを算出する工程と、

前記第 1 特徴ベクトルを、算出された前記重みによって重み付けされた前記 N 個の d 次元ベクトルの前記少なくとも一部の加重和と集約することにより前記第 2 特徴ベクトルを生成する工程と、を含む請求項 9 に記載の方法。

【請求項 12】

前記マルチモーダル統計モデルは、第 1 タスク埋め込みおよび第 2 タスク埋め込みをさら

50

に備え、前記訓練工程は、前記第 1 タスク埋め込みおよび前記第 2 タスク埋め込みのパラメータの値の推定を、前記予測子のパラメータの値の推定と共同とする工程をさらに含む、請求項 1 に記載の方法。

【請求項 13】

前記第 1 エンコーダはニューラルネットワークを含む、請求項 1 に記載の方法。

【請求項 14】

前記ニューラルネットワークは畳み込みニューラルネットワークである、請求項 13 に記載の方法。

【請求項 15】

前記ニューラルネットワークは回帰型ニューラルネットワークである、請求項 13 に記載の方法。 10

【請求項 16】

前記第 1 モダリティの前記訓練データは画像を含む、請求項 1 に記載の方法。

【請求項 17】

前記第 2 モダリティの訓練データはテキストを含む、請求項 16 に記載の方法。

【請求項 18】

1 つ以上のコンピュータハードウェアプロセッサと、
1 つ以上の非一時的なコンピュータ可読記憶媒体と、を備えるシステムであって、前記非一時的なコンピュータ可読記憶媒体は、前記 1 つ以上のコンピュータハードウェアプロセッサによって実行された場合、前記 1 つ以上のコンピュータハードウェアプロセッサに、
第 1 モダリティからの入力データおよび前記第 1 モダリティとは異なる第 2 モダリティからの入力データを含む複数のモダリティからの入力データを受信するように構成されたマルチモーダル統計モデルを訓練する方法、を実行させるプロセッサ実行可能な命令を記憶し、前記方法は 20

前記第 1 モダリティの訓練データおよび前記第 2 モダリティの訓練データを含む訓練データにアクセスする工程と、

前記マルチモーダル統計モデルを 2 段階で訓練する、訓練工程であって、前記マルチモーダル統計モデルは、前記第 1 モダリティおよび前記第 2 モダリティの入力データをそれぞれ処理する第 1 エンコーダおよび第 2 エンコーダと、第 1 モダリティ埋め込みおよび第 2 モダリティ埋め込みと、共同モダリティ表現と、予測子とを含む複数の構成要素を含み、前記第 1 モダリティ埋め込みおよび前記第 2 モダリティ埋め込みは、前記共同モダリティ表現を第 1 モダリティの空間および第 2 モダリティの空間にそれぞれ投影するための複数のベクトルおよび複数の重みを含み、前記訓練工程は、 30

前記訓練データを使用して、前記複数の構成要素のパラメータの値を推定することを含む訓練工程と、

前記マルチモーダル統計モデルの前記複数の構成要素のパラメータの推定した前記値を記憶することにより、前記マルチモーダル統計モデルを指定する情報を少なくとも部分的に記憶する工程と、を含むシステム。

【請求項 19】

前記訓練工程は、 40

前記第 1 モダリティ埋め込みおよび前記第 2 モダリティ埋め込みと前記共同モダリティ表現とのパラメータの値を推定することにより、少なくとも部分的に第 1 訓練段階を実行する、第 1 訓練段階実行工程および、

前記予測子のパラメータの値を推定することにより、少なくとも部分的に第 2 訓練段階を実行する、第 2 訓練段階実行工程、を含む請求項 18 に記載のシステム。

【請求項 20】

1 つ以上の非一時的なコンピュータ可読記憶媒体であって、前記非一時的なコンピュータ可読記憶媒体は、1 つ以上のコンピュータハードウェアプロセッサによって実行された場合、前記 1 つ以上のコンピュータハードウェアプロセッサに、第 1 モダリティからの入力データおよび前記第 1 モダリティとは異なる第 2 モダリティからの入力データを含む複数 50

のモダリティからの入力データを受信するように構成されたマルチモーダル統計モデルを訓練する方法、を実行させるプロセッサ実行可能な命令を記憶し、前記方法は前記第 1 モダリティの訓練データおよび前記第 2 モダリティの訓練データを含む訓練データにアクセスする工程と、

前記マルチモーダル統計モデルを 2 段階で訓練する、訓練工程であって、前記マルチモーダル統計モデルは、前記第 1 モダリティおよび前記第 2 モダリティの入力データをそれぞれ処理する第 1 エンコーダおよび第 2 エンコーダと、第 1 モダリティ埋め込みおよび第 2 モダリティ埋め込みと、共同モダリティ表現と、予測子とを含む複数の構成要素を含み、前記第 1 モダリティ埋め込みおよび前記第 2 モダリティ埋め込みは、前記共同モダリティ表現を第 1 モダリティの空間および第 2 モダリティの空間にそれぞれ投影するための複数のベクトルおよび複数の重みを含み、前記訓練工程は、

10

前記訓練データを使用して、前記複数の構成要素のパラメータの値を推定することを含む訓練工程と、

前記マルチモーダル統計モデルの前記複数の構成要素のパラメータの推定した前記値を記憶することにより、前記マルチモーダル統計モデルを指定する情報を少なくとも部分的に記憶する工程と、を含む非一時的なコンピュータ可読記憶媒体。

20

30

40

50