

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6925655号
(P6925655)

(45) 発行日 令和3年8月25日(2021.8.25)

(24) 登録日 令和3年8月6日(2021.8.6)

(51) Int.Cl. F1
G16C 20/70 (2019.01) G16C 20/70

請求項の数 5 (全 13 頁)

<p>(21) 出願番号 特願2019-221081 (P2019-221081) (22) 出願日 令和1年12月6日(2019.12.6) (65) 公開番号 特開2020-91872 (P2020-91872A) (43) 公開日 令和2年6月11日(2020.6.11) 審査請求日 令和1年12月6日(2019.12.6) (31) 優先権主張番号 10-2018-0156028 (32) 優先日 平成30年12月6日(2018.12.6) (33) 優先権主張国・地域又は機関 韓国(KR) (31) 優先権主張番号 10-2019-0150069 (32) 優先日 令和1年11月21日(2019.11.21) (33) 優先権主張国・地域又は機関 韓国(KR)</p>	<p>(73) 特許権者 518448323 インダストリー アカデミー コーオペレーション ファウンデーション オブ セジョン ユニバーシティー 大韓民国, 05006 ソウル, クァンジン-グ, ヌンドン-ロ, 209 (74) 代理人 110000338 特許業務法人HARAKENZO WORLD PATENT & TRADEMARK (72) 発明者 ソン, キーソン 大韓民国, 06057 ソウル, ガンナム-グ, オンジュ-ロ 146-ギル, 18, 4-302</p>
---	---

最終頁に続く

(54) 【発明の名称】 物質の分析装置

(57) 【特許請求の範囲】

【請求項1】

- (a) 2以上の元素を選択し、;
- (b) 前記2以上の元素で生成可能なものと分析された複数の化合物のデータを収集し、;
- (c) 前記収集された複数の化合物のそれぞれに関するイメージまたはスペクトル形態の分析データを準備し、;
- (d) 前記複数の化合物のうち、2種以上の化合物を選択して、前記それぞれのイメージまたはスペクトルデータを所定の混合比率に合わせて混合加工したものを含めて訓練データを作り、;
- (e) 前記訓練データを使って機械学習を遂行し、;
- (f) 実際物質から得たイメージまたはスペクトル形態の分析データを前記機械学習を通じて取得したモデルを使って、2種以上の化合物に含まれている化合物および相分率を分析する、物質の分析装置。

【請求項2】

前記複数の化合物の分析データは、既に分析されている化学分析、物質分析データを利用する、請求項1に記載の物質の分析装置。

【請求項3】

前記(c)においては、実際に分析された結果を使ったり、それぞれの化合物に関する物質情報を利用してプログラムを使って該当分析イメージまたはスペクトルをシミュレー

ションして獲得する、請求項 1 に記載の物質の分析装置。

【請求項 4】

前記イメージまたはスペクトル形態の分析データは、XRD データ、XPS データ、IR データまたは透過電子顕微鏡回折パターンである、請求項 1 に記載の物質の分析装置。

【請求項 5】

前記機械学習は CNN (Convolutional Neural Network) を通じて遂行される、請求項 1 に記載の物質の分析装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は無機物および有機物などの多様な物質に含まれた元素、相 (phase)、組成などの判別および分析を効率的に遂行できるようにする方法に関する。

【背景技術】

【0002】

粉末分析技術に基づいた物質の発見過程で最も頻りに直面する状況の一つは知られていない多相化合物の識別および定量化と関連がある。

【0003】

一般的に物質に含まれた元素、相 (phase)、成分などの分析には、XRD、XPS、EDS などの多様な分析装置が使われている。ところが、前記分析装置を通じて取得した分析結果は、該当技術分野に専門的な知識と多くの経験を有する人が相当な時間をかけないと、正確な分析結果を得られない場合が多い。

【0004】

物質分析時間を短縮するために、開発されている多様な分析ソフトウェアがある。例えば、理論結晶学的理論を基盤とするアルゴリズムを通じての分析があるが、十分な正確度を提供できない。また、機械学習方法を利用することも、訓練データセットが小規模であり、理論的な知識に基づいた過度なフィーチャーエンジニアリングを使うため、既存の規則基盤の分析ソフトウェアと大差ない結果を提供する。

【0005】

また、既存の分析ソフトウェアは、該当物質を含む製造工程ラインでリアルタイムで製造されている物質に関する情報を正確に導き出せるものは殆どないと言える。

【先行技術文献】

【特許文献】

【0006】

大韓民国公開特許公報第 10 - 2018 - 0073819 号

【発明の概要】

【発明が解決しようとする課題】

【0007】

本発明の課題は、無機物および有機物などの多様な物質に含まれた元素、相 (phase) または含量などの判別と分析を、短時間内に高精度で行うことができるようにする、物質の判別および分析方法を提供するところにある。

【課題を解決するための手段】

【0008】

前記課題を解決するために本発明は、(a) 2 以上の元素を選択する段階；(b) 前記 2 以上の元素で生成可能なものと分析された複数の化合物のデータを収集する段階；(c) 前記収集された複数の化合物のそれぞれに関するイメージまたはスペクトル形態の分析データを準備する段階；(d) 前記複数の化合物のうち、2 種以上の化合物を選択して所定の混合比率で混合し、前記それぞれのイメージまたはスペクトルデータを前記所定の混合比率に合わせて混合加工したものを含めて訓練データを作る段階；(e) 前記訓練データを使って機械学習を遂行する段階；および (f) 実際物質から得たイメージまたはスペクトル形態の分析データを前記機械学習を通じて取得したモデルを使って判別および/ま

10

20

30

40

50

たは分析する段階を含む、物質の判別および分析方法を提供する。

【0009】

前記方法において、前記2以上の元素を選択する段階において、好ましい元素の数は2～10であり、より好ましい元素の数は3～8である。

【0010】

前記方法において、前記複数の化合物のデータは現在まで遂行されている化学分析、物質分析データを利用することができる。例えば、無機物の場合、ICSDに登録された化合物を使うことができる。

【0011】

前記方法において、前記複数の化合物に関してそれぞれのイメージまたはスペクトル形態の分析データを準備する段階は、実際に分析された結果を使ったり、それぞれの化合物に関する物質情報を利用してプログラムを使って該当分析イメージまたはスペクトルを計算したのを使ってよい。

10

【0012】

前記方法において、前記所定の混合比率は例えば、A化合物、B化合物、C化合物の3種の化合物が存在する場合、前記3種の化合物の含量が質量%で0.5%単位で変更されるようにし、それに合わせてイメージまたはスペクトルデータを加工する形態であり得るが、分析する物質の特性および備えられるデータの事情に合わせて調節され得、必ずしもこれに制限されるものではない。

【0013】

20

前記方法において、前記イメージまたはスペクトル形態の分析データは機械学習が可能な形態であれば特に制限されず、例えば、XRDデータ、XPSデータ、IRデータ、透過電子顕微鏡回折パターンなどが使われ得る。

【0014】

前記方法において、前記機械学習は多様な形態のアルゴリズムが適用され得、例えばCNN(Convolutional Neural Network)を通じて遂行され得る。

【発明の効果】

【0015】

本発明によると、任意の元素からなる成分システムを設定し、このシステムで可能であると知られているすべての化合物の分析データを利用して、前記すべての化合物の混合比率をあらかじめ想定してイメージまたはスペクトルデータを加工して訓練データセットで機械学習をした後、機械学習の結果物である予測モデルを通じて、実際物質の判別および/または分析が行われるので、短時間内に正確な分析結果を導き出すことができる。例えば、鉄鋼工場ラインで、XRDデータ結果を本発明で取得したフィッティング結果を使う場合、リアルタイムで相(phase)分析が可能であるため不良発生の有無を確認できるようになる。

30

【0016】

また、本発明に係る物質の判別および/または分析方法は、従来開発された理論基盤のソフトウェア方式などと比べ、多様な成分が混合されている実際物質の相分析および相率の分析を高精度で遂行できるようにする。

40

【図面の簡単な説明】

【0017】

【図1】本発明の実施例に係る物質の判別方法に関するフローチャートである。

【図2】本発明の実施例で使われた化合物を示したものである。

【図3】本発明の一実施形態により、ICSD結晶構造データからXRD構成パラメータを利用して特定化合物のXRDデータを計算して導き出す段階を説明する図面である。

【図4】分析しようとする成分を含む無機化合物である Al_2O_3 、 Li_2O 、 SrO および $SrAl_2O_4$ の実際に分析したXRDパターンと、前記方法を通じてシミュレーションされたXRDパターンとを比較したものである。

50

【図5】各化合物の混合データを導き出す過程を示したものである。

【図6】図5で導き出された混合比率に合わせてXRDデータを加工する過程を示したものである。

【図7】混合XRDデータを導き出す具体的な条件を示したものである。

【図8】本発明の実施例でディープランニングに使った第1CNNアーキテクチャーを示したものである(aはCNN2、bはCNN3)。

【図9】本発明の実施例でディープランニングに使った第2CNNアーキテクチャーを示したものである。

【図10】「Dataset__800k__org」を使った場合、学習コストおよび正確度を示したものである。

【図11】「Dataset__800k__rand」および「Dataset__180k__rand」を使った場合、学習コストおよび正確度を示したものである。

【発明を実施するための形態】

【0018】

以下、本発明の実施例について添付された図面を参照してその構成および作用を説明する。下記の本発明の説明において、関連した公知の機能または構成についての具体的な説明が本発明の要旨を不要に曖昧にさせ得ると判断される場合にはその詳細な説明を省略する。また、ある部分がある構成要素を「含む」とする時、これは特に反対の記載がない限り、他の構成要素を除くものではなく他の構成要素をさらに含むことができることを意味する。

【0019】

図1は、本発明の実施例に係る物質の判別方法に関するフローチャートである。図1に図示されたように、本発明の実施例に係る物質の判別方法は、分析しようとする物質を構成する元素を選定する段階と、選定された元素で可能な化合物に関する情報を抽出する段階と、抽出された各化合物に対するXRDデータをシミュレーションする段階と、シミュレーションされたXRDデータを所定比率で混合して機械学習用ビッグデータを製造する段階と、CNNアーキテクチャーを通じて機械学習を遂行する段階を含んで構成される。

【0020】

[元素の選定および化合物データの収集]

本発明の実施例では、分析しようとする物質に含まれる元素として、無機発光物質の典型的な成分であるストロンチウム(Sr)、アルミニウム(Al)、リチウム(Li)および酸素(O)を選定した。

【0021】

次いで、前記4つの元素を含む化合物情報を無機結晶構造データ(ICSD、Inorganic Crystal Structure Data)を通じて確認した結果、総174個が登録されており、このうち、5個は誤ったものであるため除外した。ここに2018年ICSDバージョンには含まれていないが、最近発見された4元系化合物であってLED用発光材料で有望なSr₂LiAlO₄を追加した。

【0022】

このような過程を通じて、ストロンチウム(Sr)、アルミニウム(Al)、リチウム(Li)および酸素(O)を含む総170個の化合物情報を抽出したし、抽出された無機化合物を図2に示した。

【0023】

本発明の実施例では分析しようとする物質を無機物に選定してICSDデータを利用したが、異なる形態の物質の場合、異なる形態のデータを使用できることは言うまでもない。

【0024】

[化合物分析のためのXRDデータシミュレーション]

以上のような化合物の分析には多様な分析方法が使われ得るため、化合物の分析に適用しようとする分析方法を選択し、関連データの確保が必要である。物質の分析にはXRD

10

20

30

40

50

データ、XPSデータ、またはIRデータなどの多様なデータが使われ得るが、本発明の実施例では、無機物の相 (p h a s e) を判別する分析に一般的に使われているXRDを物質判別用分析データとして選択した。

【0025】

次に、170個の化合物に関するXRDデータがある場合、該当XRDデータを使ってもよいが、本発明の実施例では十分なXRDデータがなくても、機械学習に十分な効果を具現できるように、シミュレーション方法を通じてXRDデータを準備した。

【0026】

図3は、ICSD結晶構造データから特定化合物のXRDデータをシミュレーションする過程を示したものである。

10

【0027】

本発明の実施例ではバックグラウンドシミュレーションのための総6個の無作為パラメータを選択し、ピーク状のシミュレーションのためには総5個の無作為パラメータを使用し、ランダムノイズデータを付与方法を通じて、実際と非常に類似するシミュレーションされたXRDパターンデータを導き出した。この時、ICSD結晶情報からXRDパターンを導き出す時に格子定数を無作為で変形させてより多くの数のXRDパターンを導き出せるようにした。

【0028】

シミュレーションされたXRDパターンは、ICSDから得た構造ファクター (s t r u c t u r a l f a c t o r) および熱的ファクター (t h e r m a l f a c t o r) と、多重度 (m u l t i p l i c i t y) 、ローレンツポラリゼーションファクター (L o r e n t z a n d p o l a r i z a t i o n f a c t o r) 、吸収度 (a b s o r p t i o n) および優先方向 (p r e f e r r e d o r i e n t a t i o n) を固定変数に使用し、同時にピークプロファイル (C a g l i o t i およびミキシング媒介変数) 、背景およびホワイトノイズのような調整可能な媒介変数を使って得た。

20

【0029】

具体的には、ICSDに提示されたシンメトリー (s y m m e t r y) データから多重度を得、入射ビームに黒鉛モノクロメーターが装着されたブラッグブレンターノ (B r a g g - B r a n t a n o) ジオメトリーを使って一般実験用XRDにポラリゼーション補正が適用されたし、優先配向は存在しないと見なした。また、実験室規模のX線回折計によく現れるパラメーター値を参照して調整可能な変数 (p a r a m e t e r) を無作為で割り当てた。

30

【0030】

図4は、分析しようとする成分を含む無機化合物である Al_2O_3 、 Li_2O 、 SrO および $SrAl_2O_4$ の実際に分析したXRDパターンと、前記方法を通じてシミュレーションされたXRDパターンとを比較したものである。図4で上に表示されたパターン (緑色) は実際の実験を通じて分析された結果を示したものであり、下に表示されたパターン (茶色) は前記方法でシミュレーションされたXRDパターンを示したものである。図4に示したように、シミュレーションされたXRDパターンと実際のXRDパターン間には差を区分することが難しいほど非常に類似する形態を示した。これを通じて、本発明の実施例によりシミュレーションされたXRDパターンは機械学習用訓練データに適用できると思われ、これはシミュレーションXRDパターンを利用した機械学習を通じて樹立したモデルを使って実際物質の正確な判別につながる結果を通じても確認される。

40

【0031】

以上のような過程を通じて、総170個の化合物に対する多様なシミュレーションXRDパターンを導き出した。

【0032】

[訓練データセット作成]

図5は、各化合物の混合データを導き出す過程を示したものであり、図6は図5で導き出された混合比率に合わせてXRDデータを加工する過程を示したものであり、図7は混

50

合データを導き出す具体的な条件を示したものである。

【0033】

図5に図示されたように、A化合物、B化合物、C化合物を利用した1種(unary)、2種(binary)、3種(ternary)の混合物に対する組成の選択と混合比率の割当を通じて多様な成分と混合比率を有する混合物が導き出されるようにした。

【0034】

2種および3種の混合は、図2に示したグループに基づいて、170個ではなく38個のクラスで構成された組み合わせに基づいて生成した。その後、クラスを選択するたびに各構成要素に170個の変形物のうち一つを無作為で割り当てた。例えば、 Al_2O_3 が関心混合物の成分として選択された時、74個の変形のうち一つを Al_2O_3 に無作為で割り当てた。これを通じて Al_2O_3 は74個の変形が少なくとも一度は現れるようにした。また、 ${}_{38}C_3 \times 21 + {}_{38}C_2 \times 9 + {}_{38}C_1 (183, 521)$ の異なる混合物に対して、3種混合物は21個の固定された組成を有し、2種混合物は9個の固定された組成を有する。すなわち、混合物当たり平均4倍反復を招き得るが、多くの変形がある一部の混合物についてのみ反復的な選択を採択したため、全体データセットから同じ混合物(すなわち、同一成分および同一分率)をほとんど見出すことができなかった。このような方式でデータの準備をしないと、38個のクラスのそれぞれが同じ選択機会を有するため、例えば変形相がない Li_2O_2 に比べて74個の変形相がある Al_2O_3 の各相が選択される可能性はるかに少なくなる問題が発生する可能性がある。

【0035】

以上のような過程を通じて混合比率が導き出されると、図6に図示された方法を通じて、混合物のXRDパターンを導き出す。図6に図示された通り、A化合物でシミュレーションされたXRDパターンとB化合物でシミュレーションされたXRDパターンとC化合物でシミュレーションされたXRDパターンを、混合物の混合比率に合わせてXRDパターンを足す方法を通じて、混合物のXRDパターンを導き出した。

【0036】

混合物のシミュレーションされたXRDパターンは、2種または3種混合物組み合わせを構成するために、170個のエントリーから2~3個の化合物を選択するたびに、ピークプロファイルが無作為で選択されるようにして総800,942個のXRDパターンの間に同じものが存在しないようにした。この過程で使われた具体的な条件は図7に示した。この時、同一クラス(class)に属するエントリーで構成された2種または3種混合物組み合わせはすべて除外した。また、任意に選択された背景および白色雑音が各混合物に適用された。

【0037】

また、導き出されたXRDパターンは、無作為で訓練用データセット(600,942パターン)、有効性検査用データセット(100,000パターン)およびテストデータセット(100,000パターン)にそれぞれ分割された。

【0038】

また、データ類似性がないようにランダム組成選択および反復選択がない条件で、前記した方法に比べて縮小されたデータセットを別途に準備したし、この縮小されたデータセットの混合物数は ${}_{38}C_3 \times 21 + {}_{38}C_2 \times 9 + {}_{38}C_1 (183, 521)$ に減った。

【0039】

以上のような過程を通じて3種類のデータセットが準備された。図5の上を示したように、A化合物、B化合物、C化合物の3種21個、2種9個を等間隔で分け、図2に示した物質のうちICSD個数が多く含まれている混合物は図7に示した条件により反復して生成した800,942個のXRDパターンデータを「Dataset__800k_or_g」という。

【0040】

図5の下に示したように、A化合物、B化合物、C化合物の3種21個、2種9個をラ

10

20

30

40

50

ランダムに分けて、図2に示した物質のうちICSD個数が多く含まれている混合物は図7と同じ条件により反復して生成した800, 942個のXRDパターンデータを「Dataset_800k_Rand」という。

【0041】

A化合物、B化合物、C化合物の3種21個、2種9個をランダムに分け、反復回数なしに生成された183, 521個のXRDパターンデータを「Dataset_180k_Rand」という。

【0042】

[CNNを適用したディープランニング]

本発明の実施例ではデータセットの大きさが大きく、多相XRDパターンを扱うため、この分析に適合したCNNを適用して機械学習を遂行した。

【0043】

図8および図9は、本発明の実施例でディープランニングに使った二つのCNNアーキテクチャーを示したものである。

【0044】

図8および図9に示したように、CNNアーキテクチャーは、複数のコンボリューションレイヤ(convolution layer)とプーリングレイヤ(pooling layer)と3個の連続的に連結された全結合層(fully connected layer)で構成される。

【0045】

具体的には、図8に図示されたように、2個のコンボリューションレイヤがあるアーキテクチャーを「CNN_2」といい、図9に図示されたように、3個のコンボリューションレイヤがあるアーキテクチャーを「CNN_3」という。この二つのアーキテクチャーにおいて、フィルタ数、カーネルの大きさ、プーリング大きさ、ストライドの構成は図8および図9に図示された通りである。

【0046】

一方、図8に示されたように、本発明の実施例では、ストライドがコンボリューションレイヤのうち一つのプーリング大きさより広いものを採択したが、このような構成は既存のプーリングより顕著に早い収束を提供した。一方、CNN2およびCNN3のプーリング大きさと同じストライドでもテストした結果、収束は相対的に遅いが正確度は実質的に同じ水準と示された。

【0047】

また、コンボリューションレイヤのための活性化関数と全結合層のための線形関数として整流線形ユニット(ReLU)が採択されたし、ドロップアウトは全結合層に対してのみ具現された。全結合層に対する最終活性化機能はシグモイド機能であり、それによるコスト(または損失)機能は交差エントロピー機能であった。入力は4501×1ベクトルを、出力は38×1ベクトルにした。また、Adamオプティマイザーを使った。相の識別のためのすべてのエポック(epoch)に対して実行速度は0.001に固定された。

【0048】

以上のようなCNNアーキテクチャーで訓練データセットを使って予測モデルを導き出した。

【0049】

一方、本発明の実施例で適用したCNNアーキテクチャーは多様に変形され得、本実施例で提示したアーキテクチャーに制限されない。さらに、必ずしもCNNのみを利用する必要もなく、例えば循環神経網(Recurrent Neural Network、RNN)を使ってもよく、特にシミュレーションで導き出されたXRDパターンの実測XRDパターンとの差を補完するためには、生成的敵対神経網(Generative Adversarial Network、GAN)技術が適用されてもよい。

【0050】

10

20

30

40

50

[CNN予測モデルを利用した相判別結果]

以上で説明した訓練データセットを使ってCNN2およびCNN3アーキテクチャーを通じての学習結果を通じて導き出された予測モデルを使用して、相判別 (phase identification) が可能であるかをテストした。

【 0051 】

(1) Dataset__800k__org

図10は「Dataset__800k__org」を使った場合、学習コストおよび正確度を示したものである。図9で確認されるように、CNN2アーキテクチャーやCNN3アーキテクチャーはいずれも有効性検査結果の正確度がほとんど100%に到達したし、CNN2およびCNN3に対して有効性検証コストがそれぞれ0.007および0.0018に減少した。

10

【 0052 】

「Dataset__800k__org」を使って訓練したCNN2およびCNN3のテストのために、訓練データセットと重ならないように100,000個のシミュレーションXRDパターンを準備したし、また、3種の化合物 $Li_2O - SrO - Al_2O_3$ と $SrAl_2O_4 - SrO - Al_2O_3$ を多様な分率で乾式混合してそれぞれ50個の混合物サンプルを製造して、実際のXRDパターンを獲得して相判別テストに使った。

【 0053 】

下記の表1は前記二つのテストデータセットを使ってテストした結果を示したものである。

20

【 0054 】

【表1】

CNNアーキテクチャー		Dataset__80K__org	
		CNN2	CNN3
相判別 (2 Epochs)	シミュレーションXRD テストデータセット		
	100,000パターン	99.60%	100%
	実際のXRD テストデータセット		
	$Li_2O - SrO - Al_2O_3$ (50パターン)	100%	100%
	$SrAl_2O_4 - SrO - Al_2O_3$ (50パターン)	97.33%	98.67%

30

【 0055 】

前記表1に示されたように、シミュレーションされたテストデータセットはCNN2およびCNN3アーキテクチャーにいずれもテスト正確度がほぼ完璧な水準と示された。また、 $Li_2O - SrO - Al_2O_3$ データセットのテスト結果もCNN2およびCNN3のすべてに対して正確度が100%であって完璧に一致した。

40

【 0056 】

一方、 $SrAl_2O_4 - SrO - Al_2O_3$ データセットの場合、テスト正確度が多少低下する結果を示すが、これは混合原料として使われた $SrAl_2O_4$ (SAO) 粉末に一定量の不純物相が含まれたことによるものと確認されたし、これは若干の誤予測された結果がCNNモデルのエラーではないという点を意味する。

【 0057 】

このようなテスト結果から、本発明の実施例に係る方法が、多様な物質が混合された多成分系物質に含まれた相 (phase) の判別において、非常に高い正確度を提供できる

50

ということが分かる。

【0058】

(2) Dataset_800k_rand / Dataset_180k_rand による相識別

データセット「Dataset_800k_rand」と「Dataset_180k_rand」は前記データセットとは異なって、2種および3種混合物の設定において任意的な方法を使ったものであって、訓練データセットを準備する時、混合物データを生成する時にブレンド方法の差による影響を比較するためのものである。

【0059】

図11に示されたように、Dataset_800k_orgを使ってCNNを遂行した場合と実質的に同じコストおよび正確度が達成された。

10

【0060】

下記の表2は前記したテストデータセットを利用してテストした結果を示したものである。

【0061】

【表2】

CNNアーキテクチャー		Dataset_80K	Dataset_180K
		_rand	_rand
		CNN3	CNN3
相判別 (Epochs)	シミュレーションXRD テストデータセット		
	100,000パターン	100%	(23,000パターン) 99.76%
	実際のXRD テストデータセット		
	Li ₂ O_SrO_Al ₂ O ₃ (50パターン)	100%	98.67%
	SrAl ₂ O ₄ _SrO_ Al ₂ O ₃ (50パターン)	98%	97.33%

20

30

【0062】

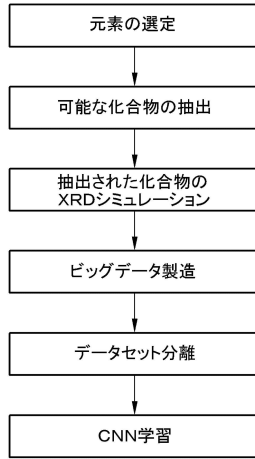
表2のように、シミュレーションされたテストデータセットと実際のLi₂O-SrO-Al₂O₃テストデータセットは、「Dataset_800k_rand」で訓練する時にほぼ100%の正確度を示した。一方、「Dataset_180k_rand」で訓練すると、シミュレーションされたテストデータセットと実際のテストデータセットでいずれもテスト正確度が多少低下する傾向を示すが、これは訓練データセットの大きさが相対的に小さいことに起因すると思われる。

40

【0063】

以上のような結果から、訓練データセットを準備する時に混合物のブレンド方法の差による影響は殆どなく、データセットの大きさにより正確度に若干の差が存在し得ることが分かる。

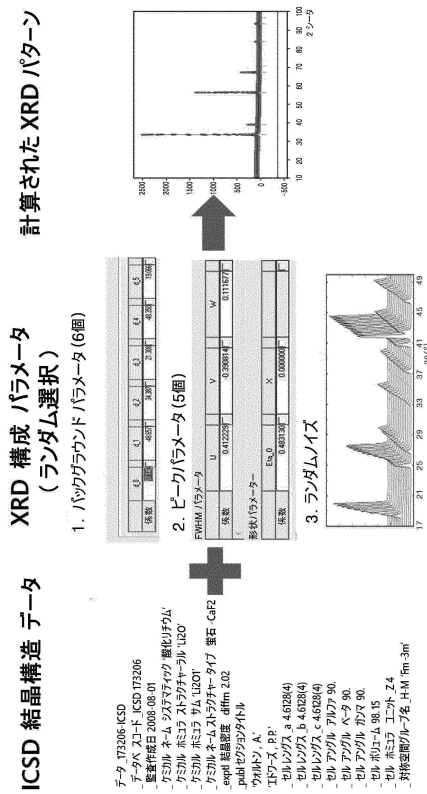
【 図 1 】



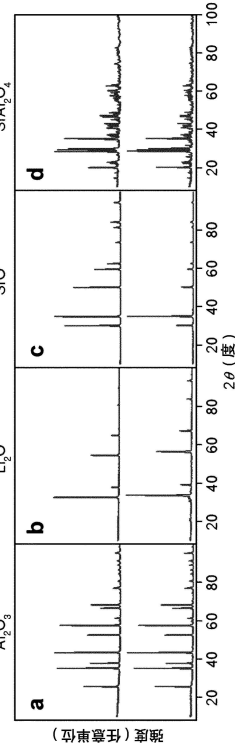
【 図 2 】

タイプ	個数	グループ 数字
コランダム-Al2O3 (Al-O)	74	0
スピネル-欠陥 (Al-O)	9	1
AlFeO3 (Al-O)	3	2
Ga2O3 (Al-O)	2	3
ガンマ (Al-O)	2	4
Ag3bO3 (Al-O)	1	5
Cr2Mg4O4 (Al-O)	1	6
蛍石-CaF2 (Al-O)	1	7
NaCl (Al-O)	1	8
Rh2S3 (Al-O)	1	9
無所属 (Al-O)(23560)	1	10
無所属 (Al-O)(40200)	1	11
蛍石-CaF2, Li2O (Li-O)	14	12
Li2O (Li-O)	1	13
無所属 (Li-O)(24143)	1	14
無所属 (Li-O)(108886)	1	15
NaCl (Sr-O)	7	16
CaC2 (Sr-O)	6	17
LiAlO2-ガンマ (Li-Al-O)	7	18
Li5AlO4 (Li-Al-O)	2	19
スピネル-LiFe5O8 (Li-Al-O)	2	20
スピネル-欠陥 (Li-Al-O)	1	21
デラフォサイト-NaCrS2 (Li-Al-O)	1	22
LiFeO2-アルファ (Li-Al-O)	1	23
Li5GaO4-アルファ (Li-Al-O)	1	24
SrAl2O4 (Sr-Al-O)	5	25
M-type-フェライト CaAl12O19 (Sr-Al-O)	4	26
Sr4Al14O25 (Sr-Al-O)	3	27
BaMgSiO4 (Sr-Al-O)	3	28
BaAl2O4 (Sr-Al-O)	2	29
Ca9Al6O18 (Sr-Al-O)	2	30
CaAlO7 (Sr-Al-O)	2	31
無所属 (Sr-Al-O)(57177.200671)	2	32
GaAlB3O7 (Sr-Al-O)	1	33
Sr5Al11O17 (Sr-Al-O)	1	34
Sr10Ga6O19 (Sr-Al-O)	1	35
無所属 (Sr-Al-O)	1	36
Sr2LiAlO4	1	37

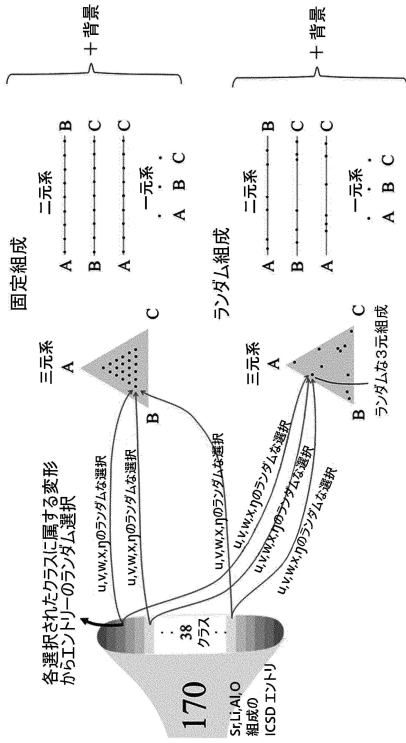
【 図 3 】



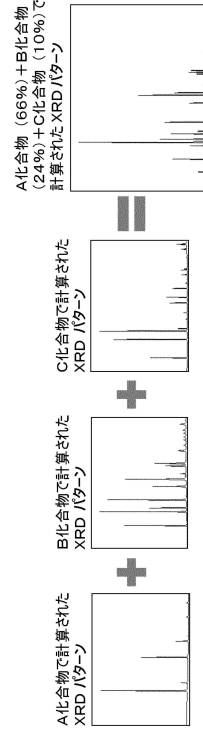
【 図 4 】



【 図 5 】



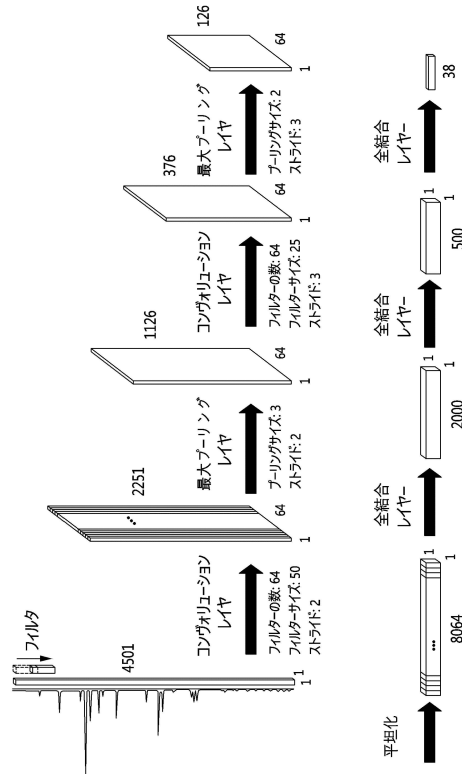
【 図 6 】



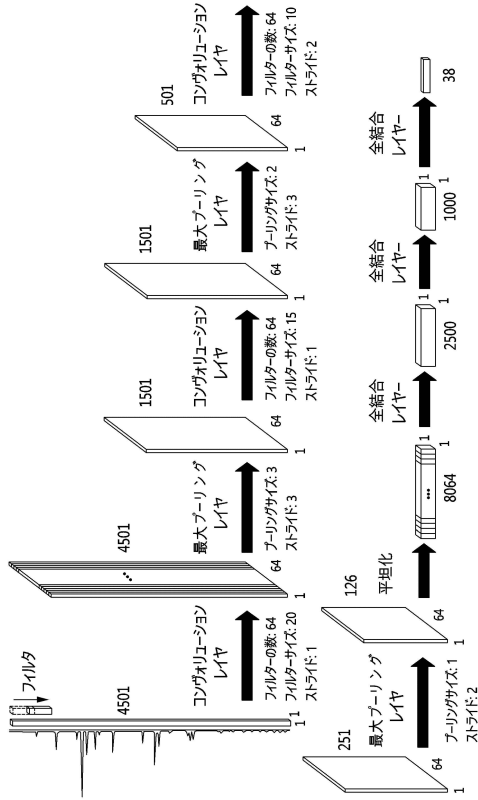
【 図 7 】

範囲	データ個数(範囲)	三元系 個数	三元系 反復回数	合計
86以上 97以下	49	21	32	32928
76以上 85以下	617	21	16	207312
10以上 75以下	2205	21	8	370440
3以上 9以下	5565	21	1	116865
				727545
範囲	データ個数(範囲)	二元系 個数	二元系 反復回数	合計
二元系は範囲なし	703	9	11	69597
範囲	データ個数(範囲)	一元系 個数	一元系 反復回数	合計
一元系は範囲なし	38	1	100	3800
				800942

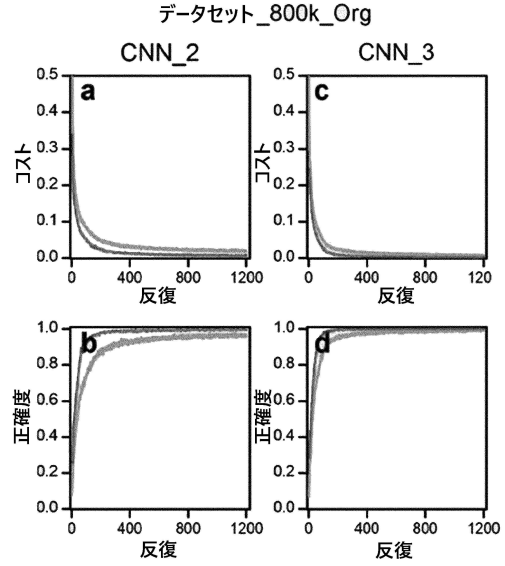
【 図 8 】



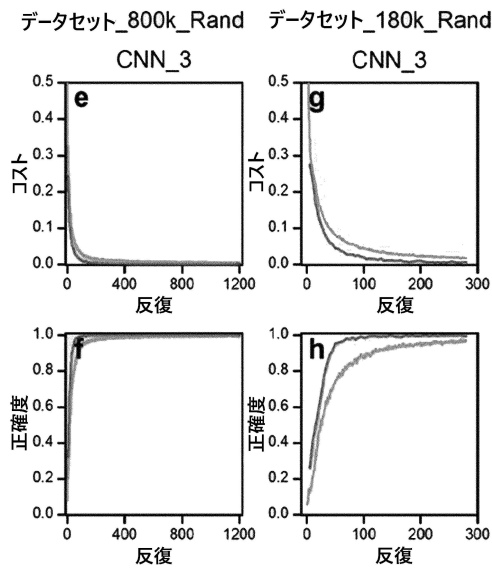
【 図 9 】



【 図 10 】



【 図 11 】



フロントページの続き

(72)発明者 イ,ジン - ウン

大韓民国, 4 8 0 0 0 プサン, ヘウンデ - グ, ウンボン - ギル, 7 7 - 7

審査官 鈴木 和樹

(56)参考文献 特開2007 - 003532 (JP, A)

特表2010 - 520471 (JP, A)

特表2010 - 527017 (JP, A)

国際公開第2018 / 025618 (WO, A1)

Felipe Oviedo、ほか10名、Fast classification of small X-ray diffraction datasets using physics-based data augmentation and deep neural networks, arXiv [ONLINE], 2018年11月20日, p. 1 - 11, [検索日: 2021/2/4], arXiv:1811.08425v1, URL, <https://arxiv.org/abs/1811.08425v1>

(58)調査した分野(Int.Cl., DB名)

G16C 10/00 - 99/00

G06Q 10/00 - 99/00

G01N 21/00 - 21/74

G01N 23/00 - 23/2276