

明 細 書

発明の名称：

情報処理装置、スクリーニング装置、情報処理方法、スクリーニング方法、及びプログラム

技術分野

[0001] 本発明は、情報処理装置、スクリーニング装置、情報処理方法、スクリーニング方法、及びプログラムに関する。

本願は、2019年5月31日に出願された日本国特許出願2019-103293号に基づき優先権を主張し、その内容をここに援用する。

背景技術

[0002] 細胞や組織の顕微鏡画像を機械学習により判定する技術が検討されている。例えば、非特許文献1には、培養細胞の顕微鏡画像を機械学習させたモデルにより、核、細胞の生死、細胞種（神経細胞であるか否か）の識別ができたことが記載されている。非特許文献2には、肺癌の病理組織の顕微鏡画像を機械学習させたモデルにより、肺腺癌、扁平上皮癌、健常な肺組織のいずれであるかの識別ができたことが記載されている。

先行技術文献

非特許文献

[0003] 非特許文献1：Christiansen E. M., et al., In Silico Labeling: Predicting Fluorescent Labels in Unlabeled Images., Cell, 173 (3), 792-803, 2018.

非特許文献2：Coudray N. et al., Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning., Nat Med., 24 (10), 1559-1567, 2018.

発明の概要

発明が解決しようとする課題

[0004] しかしながら、非特許文献 1、2 に記載された学習モデルは、細胞や組織の現在の状態を判定するものであり、被験者が将来発症するであろう疾患について予測していなかった。

[0005] 本発明は、被験者由来の多能性幹細胞から分化した細胞の画像に基づいて、被験者の将来の疾患を精度よく予測することができる情報処理装置、スクリーニング装置、情報処理方法、スクリーニング方法、及びプログラムを提供することを目的としている。

課題を解決するための手段

[0006] 本発明の一態様は、被験者由来の多能性幹細胞から分化した細胞を撮像した画像を取得する取得部と、多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像に、少なくとも前記神経変性疾患であることを示す情報が対応付けられたデータに基づいて学習されたモデルに対し、前記取得部により取得された前記画像を入力し、前記画像を入力したモデルの出力結果に基づいて、前記被験者が前記神経変性疾患を発症すること、あるいは前記神経変性疾患に対する薬剤の効果を予測する予測部と、を備える情報処理装置である。

発明の効果

[0007] 本発明の一態様によれば、被験者由来の多能性幹細胞から分化した細胞の画像に基づいて、被験者の将来の疾患を精度よく予測することができる。

図面の簡単な説明

[0008] [図1]第 1 実施形態に係る情報処理装置を含む情報処理システムの一例を示す図である。

[図2]第 1 実施形態に係る情報処理装置の構成の一例を示す図である。

[図3]第 1 実施形態に係る制御部によるランタイムの一連の処理の流れを示すフローチャートである。

[図4]クロッピング画像の生成方法を説明するための図である。

[図5]第 1 実施形態に係る予測モデルの一例を示す図である。

[図6]第 1 実施形態に係る制御部によるトレーニングの一連の処理の流れを示す図である。

すフローチャートである。

[図7]クロッピング画像の生成方法を説明するための図である。

[図8]第2実施形態に係るスクリーニング装置の構成の一例を示す図である。

[図9]実施形態の情報処理装置およびスクリーニング装置のハードウェア構成の一例を示す図である。

[図10]神経細胞を各種のコーティングを行った96ウェルプレートに播種した1日後に撮影した顕微鏡写真を示す図である。

[図11]神経細胞を各種のコーティングを行った96ウェルプレートに播種した10日後に撮影した顕微鏡写真を示す図である。

[図12]バイオアッセイの結果の一例を表す図である。

[図13]遺伝子改変技術の適用結果の一例を表す図である。

[図14]Z-VAD-FMK濃度と予測モデルのスコアとの関係性を表す図である。

[図15]提案手法を模式的に説明するための図である。

[図16]細胞のクラスタリング結果の一例を表す図である。

[図17]クラスタ0、1、及び2の一例を表す図である。

[図18]クラスタ3、4、及び5の一例を表す図である。

[図19]クラスタ6、7、及び8の一例を表す図である。

[図20]クラスタ9、10、及び11の一例を表す図である。

[図21]クラスタ12、13、及び14の一例を表す図である。

[図22]Z-VAD-FMK濃度とクラスタサイズとの関係性を表す図である。

[図23]Pathwayの解析結果の一例を表す図である。

発明を実施するための形態

[0009] 以下、本発明を適用した情報処理装置、スクリーニング装置、情報処理方法、スクリーニング方法、及びプログラムを、図面を参照して説明する。

[0010] <第1実施形態>

[全体構成]

図1は、第1実施形態に係る情報処理装置100を含む情報処理システム1の一例を示す図である。第1実施形態に係る情報処理システム1は、例えば、一以上の端末装置10と、情報処理装置100とを備える。これらの装置は、ネットワークNWを介して接続される。ネットワークNWは、例えば、インターネット、WAN (Wide Area Network)、LAN (Local Area Network)、プロバイダ端末、無線通信網、無線基地局、専用回線などを含む。図1に示す各装置の全ての組み合わせが相互に通信可能である必要はなく、ネットワークNWには、一部にローカルなネットワークが含まれてよい。

[0011] 端末装置10は、例えば、入力装置、表示装置、通信装置、記憶装置、および演算装置を備える端末装置である。具体的には、端末装置10は、パーソナルコンピュータや、携帯電話、タブレット端末などである。通信装置は、NIC (Network Interface Card) などのネットワークカード、無線通信モジュールなどを含む。例えば、端末装置10は、多能性幹細胞を用いて研究や創薬の開発などを行う施設（例えば研究機関や大学、企業）に設置されてよい。

[0012] 上述した多能性幹細胞には、例えば、胚性幹細胞 (ES細胞) や、人工多能性幹細胞 (iPS細胞)、核移植により得られるクローン胚由来の胚性幹 (ntES) 細胞、精子幹細胞 (「GS細胞」)、胚性生殖細胞 (「EG細胞」)、人工多能性幹 (iPS) 細胞などが含まれる。好ましい多能性幹細胞は、ES細胞、iPS細胞およびntES細胞である。より好ましい多能性幹細胞は、ヒト多能性幹細胞であり、特に好ましくはヒトES細胞、およびヒトiPS細胞である。さらに、本発明で使用可能な細胞は、多能性幹細胞だけでなく、多能性幹細胞を経ることなく直接所望の細胞に分化誘導させた、いわゆる「ダイレクトリプログラミング」により誘導された細胞群であってもよい。

[0013] 例えば、施設で働く従業員などは、多能性幹細胞から分化誘導させた所望の細胞を顕微鏡などで撮像し、その撮像したデジタル画像（以下、細胞画像IMGと称する）を、端末装置10を介して情報処理装置100に送信す

る。

[0014] 情報処理装置 100 は、端末装置 10 から細胞画像 IMG を受信すると、ディープラーニングを利用して、細胞画像 IMG から、分化誘導する前の多能性幹細胞の抽出元となった被験者（ユーザ）が、将来のある時点で特定の神経変性疾患を発症することを予測する。

[0015] 多能性幹細胞から分化誘導した細胞は、例えば、神経細胞、グリア細胞、血管内皮細胞、ペリサイト、脈絡叢細胞、免疫系細胞等の神経変性疾患に関連する細胞であってよい。神経変性疾患としては、例えば、アルツハイマー病、パーキンソン病、筋萎縮性側索硬化症（ALS）、脊髄小脳変性症、前頭側頭葉変性症、レビー小体型認知症、多系統萎縮症、ハンチントン病、進行性核上性麻痺、大脳皮質基底核変性症等が挙げられる。多能性幹細胞から分化した神経変性疾患の細胞とは、多能性幹細胞から分化した細胞であって、神経変性疾患の表現型を示す細胞をいう。多能性幹細胞から分化した神経変性疾患の細胞としては、例えば、神経変性疾患患者由来の多能性幹細胞から分化した細胞、神経変性疾患を発症する遺伝子変異を導入した健常者由来の多能性幹細胞から分化した細胞、健常者由来の多能性幹細胞から分化した細胞を、神経変性疾患を発症する条件下で培養した細胞等を用いることができる。

[0016] 例えば、多能性幹細胞から分化誘導した細胞が神経細胞である場合、情報処理装置 100 は、多能性幹細胞の抽出元となった被験者が、将来のある時点でアルツハイマー病を発症することを予測する。アルツハイマー病は、脳の神経細胞にアミロイドβやタウタンパク質といった物質が蓄積するという特徴の他に、神経細胞が死ぬことで脳が委縮するという特徴がある。従って、情報処理装置 100 は、多能性幹細胞から分化誘導した神経細胞が将来のある時点で神経変性疾患の表現型を示すことを予測することで、その被験者が将来のある時点でアルツハイマー病を発症するのか否かを判定する。表現型とは、ある生物のもつ遺伝子型が形質として表現されたものであり、例えば、その生物の形態、構造、行動、生理的性質などが含まれる。神経変性疾

患の表現型としては、例えば、神経細胞におけるアミロイド β の蓄積、タウタンパク質の蓄積、細胞死等が挙げられる。以下、一例として、多能性幹細胞から分化誘導した細胞が神経細胞であるものとし、神経変性疾患がアルツハイマー病であるものとし、その疾患の表現型が細胞死であるものとして説明する。

[0017] [情報処理装置の構成]

図2は、第1実施形態に係る情報処理装置100の構成の一例を示す図である。図示のように、情報処理装置100は、例えば、通信部102と、制御部110と、記憶部130とを備える。

[0018] 通信部102は、例えば、NIC等の通信インターフェースを含む。通信部102は、ネットワークNWを介して、端末装置10などと通信する。

[0019] 制御部110は、例えば、取得部112と、画像処理部114と、予測部116と、通信制御部118と、学習部120とを備える。

[0020] 制御部110の構成要素は、例えば、CPU (Central Processing Unit) やGPU (Graphics Processing Unit) などのプロセッサが記憶部130に格納されたプログラムを実行することにより実現される。制御部110の構成要素の一部または全部は、LSI (Large Scale Integration)、ASIC (Application Specific Integrated Circuit)、またはFPGA (Field-Programmable Gate Array) などのハードウェア (回路部; circuitry) により実現されてもよいし、ソフトウェアとハードウェアの協働によって実現されてもよい。

[0021] 記憶部130は、例えば、HDD (Hard Disc Drive)、フラッシュメモリ、EEPROM (Electrically Erasable Programmable Read Only Memory)、ROM (Read Only Memory)、RAM (Random Access Memory) などの記憶装置により実現される。記憶部130には、ファームウェアやアプリケーションプログラムなどの各種プログラムの他に、モデル情報132が格納される。モデル情報132については後述する。

[0022] [ランタイムの処理フロー]

以下、第1実施形態に係る制御部110によるランタイムの一連の処理の流れをフローチャートに即して説明する。ランタイムとは、既に学習された予測モデルMDLを利用する状態である。図3は、第1実施形態に係る制御部110によるランタイムの一連の処理の流れを示すフローチャートである。本フローチャートの処理は、例えば、所定の周期で繰り返し行われてよい。

[0023] まず、取得部112は、通信部102を介して、端末装置10から、神経細胞の細胞画像IMGを取得する（ステップS100）。

[0024] 次に、画像処理部114は、取得部112によって取得された細胞画像IMGをクロッピングし、細胞画像IMGよりもサイズの小さい複数の画像（以下、クロッピング画像IMG_cと称する）を生成する（ステップS102）。

[0025] 図4は、クロッピング画像IMG_cの生成方法を説明するための図である。例えば、画像処理部114は、あるアスペクト比の領域R（クロップ領域ともいう）を細胞画像IMG上でスライドさせながら走査し、領域Rと重なる画素領域をクロッピング画像IMG_cとして切り出すことで、数百枚程度（例えば120枚程度）のクロッピング画像IMG_cを生成する。

[0026] 領域Rのアスペクト比は、例えば、128×128ピクセル、256×256ピクセル、512×512ピクセル、768×768ピクセル、1020×1020ピクセルといった比率であってよい。領域Rのアスペクト比は、後述するトレーニング時のアスペクト比であるものとする。例えば、領域Rのアスペクト比が大きく、クロッピング画像IMG_cのサイズが大きくなるほど、ディープラーニングによる計算負荷が大きくなるものの、予測精度が高くなる。すなわち、予測精度を向上させるためには、クロッピング画像IMG_cのサイズを大きくし、計算速度を向上させるためには、クロッピング画像IMG_cのサイズを小さくするのが好適である。

[0027] 図3のフローチャートの説明に戻る。次に、予測部116は、画像処理部114によって生成された複数のクロッピング画像IMG_cを、モデル情報1

32が示す予測モデルMDLに入力する（ステップS104）。

[0028] モデル情報132は、神経細胞の細胞画像から、その神経細胞の細胞死を予測するための予測モデルMDLを定義した情報（プログラムまたはデータ構造）である。予測モデルMDLは、例えば、コンボリューショナルニューラルネットワーク（Convolutional Neural Network：CNN）などの種々のニューラルネットワークによって実装される。

[0029] モデル情報132には、例えば、各ニューラルネットワークを構成する入力層、一以上の隠れ層（中間層）、出力層の其々に含まれるユニットが互いにどのように結合されるのかという結合情報や、結合されたユニット間で入出力されるデータに付与される結合係数などの各種情報が含まれる。結合情報とは、例えば、各層に含まれるユニット数や、各ユニットの結合先のユニットの種類を指定する情報、各ユニットを実装する活性化関数、隠れ層のニューロン間に設けられたゲートなどの情報を含む。ニューロンを実装する活性化関数は、例えば、正規化線形関数（ReLU関数）であってもよいし、シグモイド関数や、ステップ関数、その他の関数などであってもよい。ゲートは、例えば、活性化関数によって返される値（例えば1または0）に応じて、ユニット間で伝達されるデータを選択的に通過させたり、重み付けたりする。結合係数は、例えば、ニューラルネットワークの隠れ層において、ある層のユニットから、より深い層のユニットにデータが出力される際に、出力データに対して付与される重みを含む。結合係数は、各層の固有のバイアス成分などを含んでもよい。

[0030] 図5は、第1実施形態に係る予測モデルMDLの一例を示す図である。予測モデルMDLには、例えば、K個のモデル $WL-1 \sim WL-K$ が含まれる。各モデル WL は、神経細胞の細胞画像から切り出されたクロッピング画像が入力されると、神経細胞が細胞死することの尤もらしさを示すスコアを尤度或いは確率として出力するように予め学習された弱学習器である。例えば、モデル WL はCNNを含む。各モデル WL は、互いに並列関係にある。このように複数の弱学習器を組み合わせる一つの学習モデルを生

成する手法は、アンサンブル学習と呼ばれる。

[0031] 例えば、予測モデルMDLは、弱学習器である各モデルWLのスコアを正規化し、その正規化したスコアを出力する。スコアの正規化を、数式(1)に示す。数式(1)は、例えば、全結合層(Fully-Connected層)によって実装される。

[0032] [数1]

$$S = \frac{1}{N} \sum_{i=1}^N s_i \quad \dots (1)$$

[0033] 式中Sは、正規化されたスコアを表し、 s_i は、i番目のモデルWLのスコアを表している。スコア s_i およびSは、例えば、細胞死することの確率P1と、細胞死しないことの確率P2とのそれぞれを要素とする2次元のベクトル(=[P1, P2])である。数式(1)に示すように、予測モデルMDLは、全モデルWLのスコアの和をモデルWLの総計であるKで除算することで、スコアを正規化してよい。このようにアンサンブル学習を利用することで、後述するトレーニングにおいて利用しなかった未知(未学習)のデータに対する細胞死の予測精度を向上させることができる。

[0034] 予測モデルMDLは、図5に示すように、K個のモデルWLの組み合わせであるものとしたがこれに限られない。例えば、予測モデルMDLは、一つのモデルWLであってもよい。

[0035] 図3のフローチャートの説明に戻る。次に、予測部116は、予測モデルMDLによって出力されたスコア(正規化されたスコア)が閾値以上であるのか否かを判定する(ステップS106)。

[0036] 予測部116は、スコアが閾値以上である場合、神経細胞が細胞死すると予測し(ステップS108)、スコアが閾値未満である場合、神経細胞が細胞死しないと予測する(ステップS110)。すなわち、予測部116は、スコアが閾値以上である場合、アルツハイマー病が発症する蓋然性が高いと判定し、スコアが閾値未満である場合、アルツハイマー病が発症する蓋然性が低いと判定する。

[0037] 次に、通信制御部 118 は、予測部 116 による予測結果を、通信部 102 を介して端末装置 10 に送信する（ステップ S 112）。例えば、通信制御部 118 は、神経細胞の細胞死の有無を示す情報を送信してもよいし、アルツハイマー病の発症の蓋然性を示す情報を送信してもよい。

[0038] 例えば、神経細胞の細胞死の有無を示す情報が端末装置 10 に送信された場合、その端末装置 10 を操作するユーザは、情報処理装置 100 に送信した細胞画像 IMG にうつる神経細胞が将来のある時点で死ぬ運命にあるのか、或いは死なない運命にあるのかを知ることができる。すなわち、ユーザは、神経細胞へと分化誘導する前の多能性幹細胞の抽出元の被験者が、将来アルツハイマー病を発症するのか否かを知ることができる。

[0039] [トレーニングの処理フロー]

以下、第 1 実施形態に係る制御部 110 のトレーニングの一連の処理の流れをフローチャートに即して説明する。トレーニングとは、ランタイムに利用される予測モデル MDL を学習させる状態である。図 6 は、第 1 実施形態に係る制御部 110 によるトレーニングの一連の処理の流れを示すフローチャートである。

[0040] まず、学習部 120 は、予測モデル MDL を学習するための教師データを生成する（ステップ S 200）。例えば、教師データは、多能性幹細胞から分化誘導した神経細胞を撮像した細胞画像 IMG に対して、その神経細胞が将来のある時点で細胞死することを示す情報が教師ラベル（ターゲットともいう）として対応付けられたデータである。言い換えれば、教師データは、多能性幹細胞から分化誘導した神経細胞を撮像した細胞画像 IMG を入力データとし、細胞死のような神経変性疾患の表現型を表す情報を正解の出力データとしたときに、それら入力データと出力データとを組みにしたデータセットである。将来のある時点における細胞死は、細胞画像 IMG が撮像された時点よりも顕著な神経変性疾患の表現型を表している。

[0041] 例えば、ある被験者の多能性幹細胞を分化誘導して複数の神経細胞を作製する。このようにして作製された複数の神経細胞には、アルツハイマー病を

疾患した患者の脳内と同じ環境下となるように、細胞死を促す薬剤Aが投与される。その後、薬剤Aが投与された複数の神経細胞のうち一部の神経細胞には、細胞死を抑制する薬剤B（例えばカスパーゼ阻害薬Z-VAD-FMK； $n=48$ ）が投与され、残りの神経細胞には、細胞死を抑制せず、死んだ細胞を蛍光発色させる薬剤Cが投与される。

[0042] このように薬剤Bが投与されたグループの神経細胞と、薬剤Cが投与されたグループの神経細胞とのそれぞれについて、時間経過に合わせて時系列に画像を撮像する。薬剤Bが投与されたグループの神経細胞の細胞画像IMGには、細胞死しないことを示す情報（例えばスコア0.0）が教師ラベルとして対応付けられ、薬剤Cが投与されたグループの神経細胞の細胞画像IMGには、細胞死することを示す情報（例えばスコア1.0）が教師ラベルとして対応付けられる。

[0043] このように、教師ラベルが対応付けられた複数の細胞画像IMGが教師データとして生成されると、画像処理部114は、細胞画像IMGをクロッピングし、複数のクロッピング画像IMGを生成する（ステップS202）。

[0044] 図7は、クロッピング画像IMG_Cの生成方法を説明するための図である。例えば、画像処理部114は、ある神経細胞について時系列に撮像した細胞画像群から、1つの細胞画像IMG_kを選択する。画像処理部114は、選択した細胞画像IMG_k上で所定のアスペクト比の領域Rをスライドさせながら走査し、領域Rを重なる画素領域をクロッピング画像IMG_Cとして切り出すことで、複数のクロッピング画像IMG_{C(k)}を生成する。画像処理部114は、時系列に撮像された細胞画像群の中から、前回選択した細胞画像IMG_kと異なる細胞画像IMG_{k+1}を選択し直し、同様に複数のクロッピング画像IMG_{C(k+1)}を生成する。画像処理部114は、他の神経細胞について時系列に撮像した細胞画像群に対しても同様の画像処理を行い、複数のクロッピング画像IMG_Cを生成してよい。これらの大量のクロッピング画像IMG_Cのそれぞれには、切り出し元である細胞画像IMGの教師ラベルが対応付けられる。このように予測モデルMDLを訓練させるための画像の枚数を増加さ

せることで、予測モデルMDLの予測精度を向上させる。

- [0045] 図6のフローチャートの説明に戻る。次に、学習部120は、画像処理部114によって教師データの細胞画像IMGから生成された複数のクロッピング画像IMG_cを、訓練用のクロッピング画像Tr_IMG_cと、検証用のクロッピング画像Va_IMG_cとに分け、訓練用のクロッピング画像Tr_IMG_cを、予測モデルMDLに弱学習器として含まれるK個のモデルWLの中の第i番目のモデルWL-iに入力する(ステップS204)。第i番目のモデルWL-iのパラメータ(上述した重みやバイアス成分、畳み込み層のカーネルなど)は、任意の初期値であってよい。
- [0046] 次に、学習部120は、訓練用のクロッピング画像Tr_IMG_cを入力した第i番目のモデルWL-iから出力結果、すなわちスコアs_iを取得する(ステップS206)。
- [0047] 次に、学習部120は、第i番目のモデルWL-iから取得したスコアs_iと、訓練用のクロッピング画像Tr_IMG_cに教師ラベルとして対応付けられたスコアとの誤差(損失ともいう)を算出する(ステップS208)。
- [0048] 次に、学習部120は、誤差逆伝播等の勾配法に基づいて、誤差が小さくなるように第i番目のモデルWL-iのパラメータを決定する(ステップS210)。
- [0049] 次に、学習部120は、第i番目のモデルWL-iに対する学習を所定回数E(例えば30回程度)繰り返したか否かを判定し(ステップS212)、所定回数Eに達していない場合、S204に処理を戻し、前回の処理で学習に用いた訓練用のクロッピング画像Tr_IMG_cと同じ画像を、第i番目のモデルWL-iに入力することで、第i番目のモデルWL-iを学習することを繰り返す。この際、学習部120は、学習によって更新したパラメータを記憶部130に記憶させ、パラメータを初期化した第i番目のモデルWL-iに訓練用のクロッピング画像Tr_IMG_cを入力する。これによって、第i番目のモデルWL-iに対する学習が所定回数Eに達するまでに、互いにパラメータが異なるE個のモデルWL-iが生成される。

- [0050] 一方、学習部120は、第*i*番目のモデル $WL-i$ に対する学習を所定回数*E*に達した場合、検証用のクロッピング画像 Va_IMG_c を、*E*個の第*i*番目のモデル $WL-i$ のそれぞれに入力する（ステップS214）。
- [0051] 次に、学習部120は、*E*個の第*i*番目のモデル $WL-i$ のうち、最も予測精度の高いモデル $WL-i$ を選択する（ステップS216）。例えば、学習部120は、*E*個の第*i*番目のモデル $WL-i$ のうち、検証用のクロッピング画像 Va_IMG_c を入力したときに得られるスコア s_i と教師ラベルのスコアとの誤差が最も小さいモデル $WL-i$ を選択する。
- [0052] 次に、学習部120は、予測モデルMDLに弱学習器として含まれる*K*個のモデル WL の全てについて学習したか否かを判定し（ステップS218）、まだ*K*個のモデル WL の学習が済んでいない場合、S204に処理を戻し、訓練用のクロッピング画像 Tr_IMG_c に基づいて、第*i*+1番目のモデル $WL-(i+1)$ を学習する。
- [0053] 一方、学習部120は、*K*個のモデル WL の全てについて学習した場合、本フローチャートの処理を終了する。
- [0054] 以上説明した第1実施形態によれば、情報処理装置100が、多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像に、少なくとも神経変性疾患であることを示す情報が教師ラベルとして対応付けられた教師データに基づいて学習された予測モデルMDLに対し、被験者由来の多能性幹細胞から分化した細胞を撮像した画像を入力し、画像を入力した予測モデルMDLの出力結果に基づいて、被験者が神経変性疾患を発症することを予測するため、被験者の将来の疾患を精度よく予測することができる。
- [0055] 一般的に、将来のある時点で、神経変性疾患の表現型（例えば細胞死など）が出現する細胞の細胞画像であっても、神経変性疾患を発症していない初期の頃は、細胞画像から表現型の変化を観察することが困難である。これに対して、本実施形態では、CNNによって実装される複数のモデル WL を含む予測モデルMDLを用いるため、細胞画像における肉眼では観察することが困難な細胞の構造の微小な変化や細胞間の相対的な位置関係といった特徴

を、隠れ層において畳み込みの特徴量として計算することが期待できる。この結果、人間が細胞画像を視認しても把握できない神経変性疾患を早期に発見することができる。

[0056] 更に、上述した第1実施形態によれば、予測モデルMDLを用いて、細胞に神経変性疾患の表現型が出現することを予測するため、そのような表現型が出現すると予測した場合に、その予測時点から、実際に表現型が出現するまでの期間、その細胞を観察することができる。この結果、例えば、神経変性疾患の発症の原因を究明できることが期待される。表現型が出現すると予測された細胞に様々な薬剤を投与しながら、実際に表現型が出現するまでの期間、その細胞を観察することで、投与した薬剤のメカニズムを解明できることが期待される。例えば、表現型が出現すると予測された細胞に、神経変性疾患の発症を抑制することが期待される薬剤を投与した結果、予測結果とは異なり、表現型が出現しない、或いは出現までの時間が長くなるようなことが観察された場合、投与した薬剤が神経変性疾患の発症を抑制する効果を有するという蓋然性が高いことがわかる。

[0057] 更に、上述した第1実施形態によれば、予測モデルMDLに弱学習器として含まれる複数のモデルWLをアンサンブル学習するため、予測精度の高い予測モデルMDLを生成することができる。

[0058] <第1実施形態の変形例>

以下、第1実施形態の変形例について説明する。上述した第1実施形態において、教師データは、細胞画像に対して、神経細胞が将来のある時点で細胞死するのか、或いは細胞死しないのかを表したスコアが教師ラベルとして対応付けられたデータであるものとして説明したがこれに限られない。例えば、教師データは、細胞画像に対して、上述したスコアに加えて、更に、細胞画像を撮像した時刻（第1時刻の一例）から、細胞死すると予測された時刻（第2時刻の一例）までの時間が教師ラベルとして対応付けられたデータであってよい。学習部120は、このような教師データを用いて、細胞画像が入力されると、細胞死することの確率 P_1 と、細胞死しないことの確率 P

2と、細胞死するまでの時間 t とのそれぞれを要素とする3次元のベクトル(= [P 1, P 2, t])を出力するように予測モデルMDLを学習する。予測部116は、予測モデルMDLによって出力されたベクトルの t の要素に基づいて、被験者がアルツハイマー病を発症するまでの時間を予測する。

[0059] 教師ラベルは、将来のある時点における細胞死の有無を示すスコアや細胞死するまでの時間に限られず、神経細胞の分化前の多能性幹細胞の抽出元である被験者の属性が含まれていてもよい。属性には、例えば、性別、年齢、体重、身長、生活習慣、病気の有無、家族の病歴といった種々の情報が含まれてよい。このような教師ラベルが対応付けられた細胞画像を用いて予測モデルMDLを学習することで、更に予測精度の高い予測モデルMDLを生成することができる。この結果、ランタイム時に、細胞画像に加えて被験者の属性を予測モデルMDLに入力することで、更に被験者の将来の疾患を精度よく予測することができる。

[0060] 予測モデルMDLに弱学習器として含まれる複数のモデルWLは、CNNに限られず、例えば、中間層がLSTM (Long short-term memory) であるリカレントネットワーク (Recurrent Neural Network: RNN) を含んでもよい。具体的には、モデルWLは、RNNとCNNとを組み合わせたPredNetであってよい。PredNetは、ある時刻のフレーム画像から、その時刻よりも将来の時刻のフレーム画像を予測するニューラルネットワークである。

[0061] <第2実施形態>

以下、第2実施形態について説明する。第2実施形態では、予測モデルMDLの出力結果に基づいて、被験物質が神経変性疾患の予防剤又は治療剤であるか否かを判定するスクリーニング装置100Aについて説明する。以下、第1実施形態との相違点を中心に説明し、第1実施形態と共通する点については説明を省略する。第2実施形態の説明において、第1実施形態と同じ部分については同一符号を付して説明する。

[0062] 図8は、第2実施形態に係るスクリーニング装置100Aの構成の一例を

示す図である。図示のように、スクリーニング装置 100A は、上述した第 1 実施形態に係る情報処理装置 100 の構成を含む。具体的には、スクリーニング装置 100A は、通信部 102 と、制御部 110A と、記憶部 130 とを備える。

[0063] 第 2 実施形態に係る制御部 110A は、上述した取得部 112、画像処理部 114 と、予測部 116、通信制御部 118、および学習部 120 に加えて、更に薬剤判定部 122 を備える。

[0064] 第 2 実施形態に係る取得部 112 は、被験物質を接触させた、多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像を取得する。被験物質としては特に制限されず、例えば、天然化合物ライブラリ、合成化合物ライブラリ、既存薬ライブラリ、代謝物ライブラリ等が挙げられる。本実施形態において、神経変性疾患の予防剤とは、神経変性疾患の発症前に対象に投与することにより、神経変性疾患の発症を抑制するか又は症状を軽減することができる薬剤をいう。神経変性疾患の治療剤とは、神経変性疾患の発症後に患者に投与することにより、神経変性疾患の症状を軽減することができる薬剤をいう。

[0065] 第 2 実施形態に係る学習部 120 は、上述した第 1 実施形態と同様に、教師データに基づいて予測モデル MDL を学習する。第 2 実施形態に係る教師データは、多能性幹細胞から分化した、アルツハイマー病のような神経変性疾患の細胞を撮像した細胞画像に対して、細胞死のような表現型がスコアとして対応付けられたデータである。これによって、予測モデル MDL は、細胞画像が入力されると、細胞に神経変性疾患の表現型が出現することをスコアとして出力する。

[0066] 第 2 実施形態に係る予測部 116 は、取得部 112 によって取得された画像を、学習済みの予測モデル MDL に入力する。そして、予測部 116 は、画像を入力した予測モデル MDL の出力結果に基づいて、被験物質が投与された細胞に、神経変性疾患の表現型が出現するのか否かを予測する。

[0067] 薬剤判定部 122 は、予測部 116 の予測結果に基づいて、被験物質が神

経変性疾患の予防剤又は治療剤であるか否かを判定する。

[0068] 例えば、薬剤判定部 122 は、以下の条件（１）を満たす場合に、被験物質が神経変性疾患の予防剤又は治療剤であると判定し、条件（２）を満たす場合に、被験物質が神経変性疾患の予防剤でも治療剤でもないと判定してよい。

[0069] 条件（１）：画像が入力された予測モデルMDLによって出力されたスコアが閾値以下であり、被験物質が投与された細胞に神経変性疾患の表現型が出現しないことが予測されること。

[0070] 条件（２）：画像が入力された予測モデルMDLによって出力されたスコアが閾値以上であり、被験物質が投与された細胞に神経変性疾患の表現型が出現することが予測されること。

[0071] 以上説明した第２実施形態によれば、スクリーニング装置 100A が、被験物質を接触させた、多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像を取得し、学習済みの予測モデルMDLに対し、取得した画像を入力し、その画像を入力した予測モデルMDLの出力結果に基づいて、被験物質が投与された細胞に神経変性疾患の表現型が出現するか否かを予測する。そして、スクリーニング装置 100A は、細胞に表現型が出現するか否かの予測結果に基づいて、被験物質が神経変性疾患の予防剤又は治療剤であるか否かを判定する。この結果、多能性幹細胞から分化した細胞の画像に基づいて、神経変性疾患の表現型の出現有無を予測した上で、更に、その神経変性疾患の予防剤または治療剤となり得る新薬を効率よく発見することができる。

[0072] <ハードウェア構成>

上述した実施形態の情報処理装置 100 およびスクリーニング装置 100A は、例えば、図 9 に示すようなハードウェア構成により実現される。図 9 は、実施形態の情報処理装置 100 およびスクリーニング装置 100A のハードウェア構成の一例を示す図である。

[0073] 情報処理装置 100 は、NIC 100-1、CPU 100-2、RAM 1

00-3、ROM100-4、フラッシュメモリやHDDなどの二次記憶装置100-5、およびドライブ装置100-6が、内部バスあるいは専用通信線によって相互に接続された構成となっている。ドライブ装置100-6には、光ディスクなどの可搬型記憶媒体が装着される。二次記憶装置100-5、またはドライブ装置100-6に装着された可搬型記憶媒体に格納されたプログラムがDMAコントローラ（不図示）などによってRAM100-3に展開され、CPU100-2によって実行されることで、制御部110および110Aが実現される。制御部110または110Aが参照するプログラムは、ネットワークNWを介して他の装置からダウンロードされてもよい。

[0074] [表現例1]

上記説明した実施形態は、以下のように表現することができる。

プロセッサと、

プログラムを格納したメモリと、を備え、

前記プロセッサが前記プログラムを実行することにより、

被験者由来の多能性幹細胞から分化した細胞を撮像した画像を取得し、

多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像に、少なくとも前記神経変性疾患であることを示す情報が対応付けられたデータに基づいて学習されたモデルに対し、前記取得した画像を入力し、前記画像を入力したモデルの出力結果に基づいて、前記被験者が前記神経変性疾患を発症すること、あるいは前記神経変性疾患に対する薬剤の効果を予測する、

ように構成されている、情報処理装置。

[0075] [表現例2]

上記説明した実施形態は、以下のように表現することもできる。

プロセッサと、

プログラムを格納したメモリと、を備え、

前記プロセッサが前記プログラムを実行することにより、

被験物質を接触させた、多能性幹細胞から分化した神経変性疾患の細胞を

撮像した画像を取得し、

多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像に、少なくとも前記神経変性疾患の表現型を示す情報が対応付けられたデータに基づいて学習されたモデルに対し、前記取得した画像を入力し、前記画像を入力したモデルの出力結果に基づいて、前記多能性幹細胞から分化した神経変性疾患の細胞に、前記神経変性疾患の表現型が出現するか否かを予測し、

前記予測した結果に基づいて、前記被験物質が前記神経変性疾患の予防剤又は治療剤であるのか否かを判定する、

ように構成されている、スクリーニング装置。

[0076] 以上、本発明を実施するための形態について実施形態を用いて説明したが、本発明はこうした実施形態に何ら限定されるものではなく、本発明の要旨を逸脱しない範囲内において種々の変形及び置換を加えることができる。

実施例

[0077] [実験例 1]

(ヒト i P S 細胞の樹立)

ヒト末梢血単核球細胞に、エピソーマルベクターを用いてリプログラミング因子を導入した。リプログラミング因子としては、S O X 2、K L F 4、O C T 4、L - M Y C、L I N 2 8、ドミナントネガティブ p 5 3 を用いた。数日後、細胞を回収し、i M a t r i x (ニッピ社) をコートしたディッシュに播種した。翌日、培地を S t e m F i t A K O 3 (味の素株式会社) に交換した。その後、2 日ごとに培地を交換した。リプログラミング因子の導入から 2 0 日後、i P S 細胞のコロニーを回収して拡大培養した。

[0078] [実験例 2]

(神経細胞の作製)

直接変換技術を利用して実験例 1 で樹立した i P S 細胞を神経細胞に分化させた。テトラサイクリン誘導プロモーター (t e t O) 下のヒトニューロゲニン 2 (N G N 2) 遺伝子の c D N A を、p i g g y B a c トランスポゾンシステム及び L i p o f e c t a m i n e L T X (サーモフィッシャー

サイエンティフィック社)を用いてiPS細胞に導入した。tetO::NGN2を含むベクターを使用した。G418二硫酸塩(ナカライテスク社)の抗生物質選択の後、コロニーを選び出し、NGN2の一時的発現を誘導することによって神経細胞に効率的に分化できるサブクローンを選択した。

[0079] [実験例3]

(カスパーゼ3のライブイメージング検出)

アポトーシスを起こした神経細胞における活性化カスパーゼ3は、核酸結合色素(製品名「CellEvent Caspase-3/7 Green Detection Reagent」、サーモフィッシャーサイエンティフィック社)に結合した4アミノ酸ペプチドDEVD(配列番号1)からなる蛍光発生基質を使用することによって検出した。DEVDペプチドが蛍光色素のDNAへの結合を阻害するため、上記の蛍光発生基質は本質的に非蛍光性である。しかしながら、死細胞中の活性化カスパーゼ3がDEVDペプチドを切断すると、蛍光色素がDNAに結合して蛍光を生じ、標準的なタイムラプスイメージングシステム(製品名「Incucyte ZOOM」、エッセンバイオサイエンス社)の標準FITC-グリーンフィルターセットを用いて検出することができる(露光時間400ms)。画像は薬物投与後6時間ごとに経時的に撮影された。

[0080] [実験例4]

(教師データの準備)

タイムラプスイメージングシステム(製品名「Incucyte ZOOM」、エッセンバイオサイエンス社)を使用してキャプチャされたシングルウェル画像の元のサイズは1,392×1,038ピクセルであり、この画像から4つの画像を切り取ってクロッピング画像として抽出した。元の画像を4つの領域に均等に分割し、各領域の中央から正方形の画像を切り取った。トリミングされたクロッピング画像のサイズは、幅×高さが、128×128、256×256、512×512、768×768、1020×1,020から選択した。元の画像のサイズが1,392×1,038

38であるため、クロップサイズが519×519より大きい場合、クロップ領域Rは元の画像からはみ出す。この場合、クロップ領域Rは元の画像の端に合うように移動した。

[0081] 神経細胞にカスパーゼ3阻害剤であるZ-VAD-FMK（プロメガ社）を投与したウェルから取得した画像は「生きている」クラス又は「1」として分類した。カスパーゼ3阻害剤を投与していないウェルから取得した画像は「死」クラス又は「0」として分類した。すなわち、カスパーゼ3阻害剤を投与したウェルから取得した画像には、神経細胞が「生きている」ことを示す教師ラベルを対応付け、カスパーゼ3阻害剤を投与していないウェルから取得した画像には、神経細胞が「死んだ」ことを示す教師ラベルを対応付けた。

[0082] 実験は3枚の96ウェルプレートを用いて行った。3枚のプレートのうち、ジメチルスルホキシド（DMSO）を投与した群及びZ-VAD-FMK（プロメガ社）を投与した群の、薬剤適用の24～48時間後の2枚のプレートのA列及びL列のウェルをトレーニングデータセット（1クラスあたり480画像）に使用し、残りの列のウェルを検証データセットに使用した（1クラスあたり160画像）。残りの1枚のプレートの薬物適用後24時間の時点での画像を試験データセットとして使用した（1クラスあたり64画像）。

[0083] [実験例5]

（トレーニング、検証、試験）

予測モデルMDLの構築、トレーニング、検証及び試験について説明する。予測モデルMDLの構築、トレーニング、検証及び試験には、ディープラーニングフレームワークのPyTorch（<https://pytorch.org/>）を使用した。ニューラルネットワークのアーキテクチャとして、inception-v3モデルをモデル（弱学習器）WLとして使用した。医学的又は生物学的データのようなトレーニングサンプルの数が限られている場合に有効な微調整アプローチを採用した。ニューラルネットワー

クは ImageNet データセットで事前にトレーニングされていた。ImageNet データセットは、1,000 のカテゴリに分類されたオブジェクトの 120 万のラベル付きイメージを含んでいた。事前にトレーニングされたニューラルネットワークの最後の全結合 (FC) 層におけるノードの数はクラスの数 (= 2) に調整され、すべての層における重みはトレーニング段階中に再最適化された。トレーニング用の 256×256 ピクセルの画像は、50% の確率で水平方向にランダムに反転され、次に 224×224 ピクセルの領域がランダムに画像から切り取られた。切り取られた 224×224 ピクセルの画像は、最終的にニューラルネットワークへの入力として使用された。クロスエントロピーを損失関数として使用した。誤差はネットワークを介して逆伝播され、重みはミニバッチを用いた確率的勾配法 (Stochastic Gradient Descent; SGD) によって最適化された。SGD におけるモーメント係数は 0.9 に設定した。

[0084] 検証データ (検証用のクロッピング画像 Va_IMG_c) に対して各エポックで分類精度を計算した。エポックとは、所定回数 E に達するまでにモデル WL を繰り返す学習する際の反復処理のことである。ネットワークが最良の分類精度を示したときのエポックにおけるネットワークのパラメータを各トレーニング試行で保存した。すなわち、 E 個のパラメータを保存した。

[0085] 例えば、 $E = 20$ とし、20 回のトレーニング試行を実施した。これによって、20 個のパラメータを保存した。最初に、ミニバッチサイズ B を 16 に設定し、SGD の学習率を $lr = 0.001$ にしてから、さまざまなクロップサイズ (幅 $\times$ 高さ) での 20 回の試行の検証データに対して平均精度を評価した。クロップサイズを固定した後、検証に対する精度を評価することによって、ミニバッチサイズ $B = \{4, 8, 16, 32\}$ 及び学習率 $lr = \{0.0005, 0.001, 0.0015, 0.002\}$ にわたってグリッド検索を実行し、検証データに対する確からしさを評価した。

[0086] 予測テストでは、20 個のトレーニング済みネットワークの最後の全結合層からの平均出力 (正規化されたスコア) をテストデータの分類に使用した

。検証データ及び試験データの両方において、 224×224 ピクセルの画像を、サイズ変更された 256×256 ピクセルの画像の中心から反転せずにトリミングした。ROC (Receiver Operating Characteristic) 分析における精度、F1スコア、及び曲線下面積 (AUC) を計算し、3重交差検証によって検証した。各データセットのプレートを回転させ、3回の試行の結果の平均を計算することで、検証データに対する確からしさを評価した。

[0087] 機械学習に係る上述した手法はあくまでも一例であり、周知の他の手法を採用することができる。例えば、誤差の最小化として、SGDの代わりに、AdaGrad、RMSprop、AdaDelta、Adam (Adaptive moment estimation) などを用いてもよい。

[0088] [実験例6]

(細胞培養容器のコーティングの検討)

iPS細胞から分化した神経細胞を長期間観察するのに適した細胞培養容器のコーティングを検討した。

[0089] 神経細胞としては、実験例2で選択した細胞にNGN2の一時的発現を誘導することにより得られた神経細胞を使用した。コーティング剤としては、ポリ-D-リジン (PDL、シグマ社) 終濃度1% (v/v)、マトリゲル (コーニング社) 終濃度2% (v/v)、Synthemax 11 SC (コーニング社) 終濃度2% (v/v)、CellNest (富士フイルム社) 終濃度1% (v/v) を様々な組み合わせで使用した。比較のために、コーティングを行わなかった培養ディッシュについても同様の検討を行った。

[0090] 神経細胞を各種のコーティングを行った96ウェルプレートに播種し、1日後及び10日後の細胞の様子を顕微鏡で観察した。図10は、1日後に撮影した顕微鏡写真を示す図であり、図11は、10日後に撮影した顕微鏡写真を示す図である。

[0091] 図10及び11中、PDL (シグマ社) でコートしたことを「P」と示し、マトリゲル (コーニング社) でコートしたことを「M」と示し、Synth

hemax 11 SC (コーニング社) でコートしたことを「S」と示し、CellNest (富士フイルム社) でコートしたことを「C」と示す。PDL (シグマ社) 及びマトリゲル (コーニング社) でコートしたことを「PM」と示し、その他のコーティング剤の組み合わせにおいても同様である。PDL (シグマ社)、マトリゲル (コーニング社) 及びSynthamax 11 SC (コーニング社) でコートしたことを「PMS」と示し、その他のコーティング剤の組み合わせにおいても同様である。コーティングを行わなかったことを「Non」と示す。

[0092] その結果、PMSCコーティング、すなわち、PDL (シグマ社) 終濃度 1% (v/v)、マトリゲル (コーニング社) 終濃度 2% (v/v)、Synthamax 11 SC (コーニング社) 終濃度 2% (v/v)、及び、CellNest (富士フイルム社) 終濃度 1% (v/v) の混合物でコーティングした細胞培養容器が、最も神経細胞の長期間の観察に適していることが明らかとなった。

[0093] [実験例 7]

カスパーゼ 3 の活性化の程度を、96 ウェルプレートを用いて確認した。図 12 は、バイオアッセイの結果の一例を表す図である。プレートの A 列から H 列までの全ウェルについて、神経細胞画像上において緑色に蛍光発色した領域面積 (以下、グリーンシグナルエリアと称する) を算出し、A 列から H 列のそれぞれの列ごとにグリーンシグナルエリアの平均値を算出した。グリーンシグナルエリアは、細胞死が生じた領域の面積である。そして、時間経過に応じたグリーンシグナルエリアの平均値の変化をグラフにプロットした。この結果、列ごとに有意な差異はなく、アッセイの安定性が示された。

[0094] 図 13 は、遺伝子改変技術の適用結果の一例を表す図である。家族性アルツハイマー病は、遺伝子 PSEN1 G384A 変異が原因であることが知られている。図中の AD__clone1 及び AD__clone2 は、PSEN1 G384A 変異が生じた神経細胞画像のグリーンシグナルエリアの時間経過に応じた変化を表している。Corrected__clone1~4

は、PSEN1 G384A変異が生じた神経細胞を遺伝子改変技術を用いて変異を野生型に正常化させたときのグリーンシグナルエリアの時間経過に応じた変化を表している。図示の通り、AD__clone1及びAD__clone2に比べて、Corrected__clone1～4の方が、グリーンシグナルエリアの平均値が小さく、遺伝子改変技術によって細胞死が緩和されたことが示された。

[0095] 図14は、Z-VAD-FMK濃度と予測モデルMDLのスコアとの関係性を表す図である。図中の縦軸は、予測モデルMDLのスコアを表しており、横軸は、Z-VAD-FMK濃度を表している。スコアは、上述した通り、細胞の生死に関する確率的な数値を表している。スコアが大きいほど細胞が生存する蓋然性が高くなり、スコアが小さいほど細胞のカスパーゼ3が活性化したりアポトーシスを起こしたりする蓋然性が高くなる。図示の通り、Z-VAD-FMKの濃度を高くするほどスコアが高くなるという相関性を示した。

[0096] [新規手法の提案]

図15は、提案手法を模式的に説明するための図である。図示のように、神経細胞にZ-VAD-FMKを添加し、細胞死が生じるよりも前の段階で（カスパーゼ3の活性化が生じるよりも前の段階で）、個々の細胞の遺伝子発現の変化を観察する（図のsingle cell RNA sequencing=WHAT）。同時に、画像を取得しておき、どこでその変化が生じているかをCNNをベースとした予測モデルMDLを用いて確認する（図中のWHERE）。このようにして、どこで何が、細胞死の前に生じているのかどうか、分子を同定する方法を提案する。

[0097] [実験例8]

上記の提案手法を用いてSingle cell解析を行った。図16は、細胞のクラスタリング結果の一例を表す図である。神経細胞に添加するZ-VAD-FMKの濃度を0 [μ M]、0.2 [μ M]、25 [μ M]と変えながら、UMAP (uMAP) と呼ばれる可視化手法（次元圧縮法の一種

)を利用して細胞の集団を0～14までの計15種類のクラスタに分離した。

[0098] 図17は、クラスタ0、1、及び2の一例を表す図である。図中の1行目はクラスタ0を表しており、2行目はクラスタ1を表しており、3行目はクラスタ2を表している。1列目は、Z-VAD-FMKの濃度が0 [μ M]であることを表しており、2列目は、Z-VAD-FMKの濃度が0.2 [μ M]であることを表しており、3列目は、Z-VAD-FMKの濃度が2.5 [μ M]であることを表している。

[0099] 図18は、クラスタ3、4、及び5の一例を表す図である。図中の1行目はクラスタ3を表しており、2行目はクラスタ4を表しており、3行目はクラスタ5を表している。図19は、クラスタ6、7、及び8の一例を表す図である。図中の1行目はクラスタ6を表しており、2行目はクラスタ7を表しており、3行目はクラスタ8を表している。図20は、クラスタ9、10、及び11の一例を表す図である。図中の1行目はクラスタ9を表しており、2行目はクラスタ10を表しており、3行目はクラスタ11を表している。図21は、クラスタ12、13、及び14の一例を表す図である。図中の1行目はクラスタ12を表しており、2行目はクラスタ13を表しており、3行目はクラスタ14を表している。これらの図における列も、図17同様に、1列目がZ-VAD-FMKの濃度が0 [μ M]であることを表し、2列目がZ-VAD-FMKの濃度が0.2 [μ M]であることを表し、3列目がZ-VAD-FMKの濃度が2.5 [μ M]であることを表している。

[0100] 図16から図21におけるUMAP展開した細胞集団のクラスタ形状から見て取れるように、ほとんどの細胞集団は同一の形状であった。つまり、Z-VAD-FMK濃度は、ほとんどの細胞集団に影響を与えなかった。

[0101] 図22は、Z-VAD-FMK濃度とクラスタサイズとの関係性を表す図である。カスパーゼ3が活性化しており、全細胞死に占める割合が数[%]のクラスタであり、尚且つ、Z-VAD-FMKを添加するとともに（すなわちアポトーシスを抑制するとともに）、自身のクラスタに含まれる細胞数

が減少している、という条件を満たすクラスタを14つのクラスタの中から探した。その結果、条件を満たすものとしてクラスタ12が同定された。

[0102] クラスタ12に特異性が高く発現している遺伝子群を含むP a t h w a yを解析した。図23は、P a t h w a yの解析結果の一例を表す図である。図示の通り、Z-VAD-FMKを添加したことに伴い、細胞数が減少するクラスタ（言い換えれば細胞死に向かっているクラスタ）において、特異的に高いアポトーシス及び細胞死に関わるP a t h w a yに寄与していることが示された。これにより、本実験手法の妥当性が示された。

請求の範囲

- [請求項1] 被験者由来の多能性幹細胞から分化した細胞を撮像した画像を取得する取得部と、
- 多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像に、少なくとも前記神経変性疾患であることを示す情報が対応付けられたデータに基づいて学習されたモデルに対し、前記取得部により取得された前記画像を入力し、前記画像を入力したモデルの出力結果に基づいて、前記被験者が前記神経変性疾患を発症すること、あるいは前記神経変性疾患に対する薬剤の効果を予測する予測部と、
- を備える情報処理装置。
- [請求項2] 前記モデルの学習に利用されるデータは、前記神経変性疾患の細胞を撮像した画像に対して、前記神経変性疾患の細胞を撮像した第1時刻よりも将来の第2時刻において、前記第1時刻の時点よりも顕著な神経変性疾患の表現型を示す情報が少なくとも対応付けられたデータであり、
- 前記モデルは、前記細胞を撮像した画像が入力されると、前記細胞が前記表現型を示すことの尤もらしさを示すスコアを出力し、
- 前記予測部は、前記モデルによって出力された前記スコアに基づいて、前記被験者が前記神経変性疾患を発症することを予測する、
- 請求項1に記載の情報処理装置。
- [請求項3] 前記モデルの学習に利用されるデータは、前記神経変性疾患の細胞を撮像した画像に対して、更に、前記第1時刻から前記第2時刻までの時間が対応付けられたデータであり、
- 前記モデルは、前記細胞を撮像した画像が入力されると、前記細胞が前記表現型を示すまでの時間を出力し、
- 前記予測部は、更に、前記モデルによって出力された前記時間に基づいて、前記被験者が前記神経変性疾患を発症するまでの時間を予測する、

請求項 2 に記載の情報処理装置。

[請求項4] 前記モデルの学習に利用されるデータには、健常者由来の多能性幹細胞から分化した細胞を撮像した画像に、前記細胞が神経変性疾患の表現型を示さないという情報が対応付けられたデータが更に含まれる、

請求項 1 から 3 のうちいずれか一項に記載の情報処理装置。

[請求項5] 前記データに基づいて前記モデルを学習する学習部を更に備える、
請求項 1 から 4 のうちいずれか一項に記載の情報処理装置。

[請求項6] 前記モデルには、複数の弱学習器が含まれ、
前記学習部は、前記複数の弱学習器をアンサンブル学習する、
請求項 5 に記載の情報処理装置。

[請求項7] 前記予測部は、前記取得部により取得された前記画像を、前記複数の弱学習器のそれぞれに入力し、前記画像を入力した複数の弱学習器のそれぞれの出力結果に基づいて、前記被験者が神経変性疾患を発症することを予測する、

請求項 6 に記載の情報処理装置。

[請求項8] 前記多能性幹細胞から分化した細胞は、神経細胞、グリア細胞、血管内皮細胞、ペリサイト、脈絡叢細胞又は免疫系細胞であり、

前記神経変性疾患は、アルツハイマー病、パーキンソン病、筋萎縮性側索硬化症（ALS）、脊髄小脳変性症、前頭側頭葉変性症、レビー小体型認知症、多系統萎縮症、ハンチントン病、進行性核上性麻痺、又は大脳皮質基底核変性症である、

請求項 1 から 7 のうちいずれか一項に記載の情報処理装置。

[請求項9] 被験物質を接触させた、多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像を取得する取得部と、

多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像に、少なくとも前記神経変性疾患の表現型を示す情報が対応付けられたデータに基づいて学習されたモデルに対し、前記取得部により取得され

た前記画像を入力し、前記画像を入力したモデルの出力結果に基づいて、前記多能性幹細胞から分化した神経変性疾患の細胞に、前記神経変性疾患の表現型が出現するか否かを予測する予測部と、

前記予測部の予測結果に基づいて、前記被験物質が前記神経変性疾患の予防剤又は治療剤であるか否かを判定する判定部と、

を備えるスクリーニング装置。

[請求項10] 前記判定部は、前記多能性幹細胞から分化した神経変性疾患の細胞に前記表現型が出現しないことが前記予測部によって予測された場合、前記被験物質が前記神経変性疾患の予防剤又は治療剤であると判定する、

請求項9に記載のスクリーニング装置。

[請求項11] コンピュータが、
被験者由来の多能性幹細胞から分化した細胞を撮像した画像を取得し、

多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像に、少なくとも前記神経変性疾患であることを示す情報が対応付けられたデータに基づいて学習されたモデルに対し、前記取得した画像を入力し、前記画像を入力したモデルの出力結果に基づいて、前記被験者が前記神経変性疾患を発症すること、あるいは前記神経変性疾患に対する薬剤の効果を予測する、

情報処理方法。

[請求項12] コンピュータが、
被験物質を接触させた、多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像を取得し、

多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像に、少なくとも前記神経変性疾患の表現型を示す情報が対応付けられたデータに基づいて学習されたモデルに対し、前記取得した画像を入力し、前記画像を入力したモデルの出力結果に基づいて、前記多能性幹細胞

胞から分化した神経変性疾患の細胞に、前記神経変性疾患の表現型が出現するか否かを予測し、

前記予測した結果に基づいて、前記被験物質が前記神経変性疾患の予防剤又は治療剤であるのか否かを判定する、

スクリーニング方法。

[請求項13]

コンピュータに、

被験者由来の多能性幹細胞から分化した細胞を撮像した画像を取得することと、

多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像に、少なくとも前記神経変性疾患であることを示す情報が対応付けられたデータに基づいて学習されたモデルに対し、取得した前記画像を入力し、前記画像を入力したモデルの出力結果に基づいて、前記被験者が前記神経変性疾患を発症すること、あるいは前記神経変性疾患に対する薬剤の効果を予測することと、

を実行させるためのプログラム。

[請求項14]

コンピュータに、

被験物質を接触させた、多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像を取得することと、

多能性幹細胞から分化した神経変性疾患の細胞を撮像した画像に、少なくとも前記神経変性疾患の表現型を示す情報が対応付けられたデータに基づいて学習されたモデルに対し、前記取得した画像を入力し、前記画像を入力したモデルの出力結果に基づいて、前記多能性幹細胞から分化した神経変性疾患の細胞に、前記神経変性疾患の表現型が出現するか否かを予測することと、

前記予測した結果に基づいて、前記被験物質が前記神経変性疾患の予防剤又は治療剤であるか否かを判定することと、

を実行させるためのプログラム。