(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2022年12月22日(22.12.2022)

WIPO | PCT

(10) 国際公開番号
WO 2022/264331 A1

(51) 国際特許分類:
G06N 3/02 (2006.01)

(21) 国際出願番号： PCT/JP2021/022916

(22) 国際出願日： 2021年6月16日(16.06.2021)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(71) 出願人:三菱電機株式会社(**MITSUBISHI ELECTRIC CORPORATION**) [JP/JP]; 〒1008310 東京都千代田区丸の内二丁目７番３号 Tokyo (JP).

(72) 発明者: 小関 義博 (**KOSEKI, Yoshihiro**); 〒1008310 東京都千代田区丸の内二丁目７番３号 三菱電機株式会社内 Tokyo (JP).

(74) 代理人: 弁理士法人クロスボーダー特許事務所 (**CROSS-BORDER PATENT FIRM**); 〒2470056 神奈川県鎌倉市大船二丁目１７番１０号３階 Kanagawa (JP).

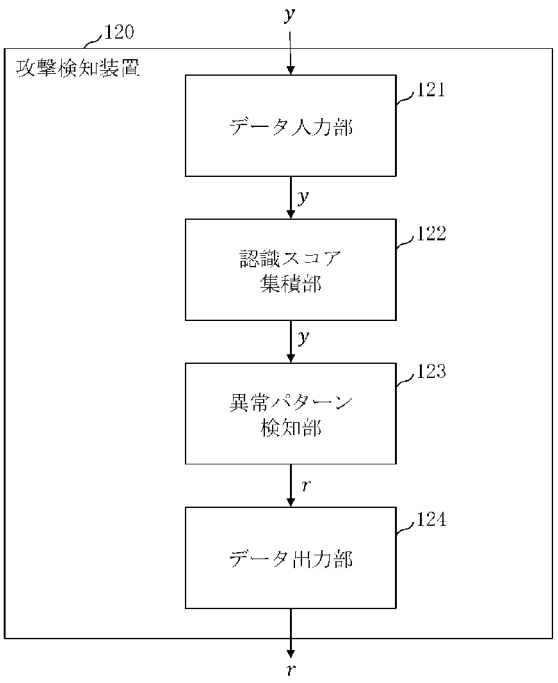
(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,

WO 2022/264331 A1

(54) **Title:** ATTACK DETECTION DEVICE, ADVERSARIAL SAMPLE PATCH DETECTION SYSTEM, ATTACK DETECTION METHOD, AND ATTACK DETECTION PROGRAM

(54) 発明の名称： 攻撃検知装置、敵対的サンプルパッチ検知システム、攻撃検知方法、及び、攻撃検知プログラム

[図3]



120... ATTACK DETECTION DEVICE
121... DATA INPUT UNIT
122... RECOGNITION SCORE ACCUMULATION UNIT
123... ABNORMAL PATTERN DETECTION UNIT
124... DATA OUTPUT UNIT

(57) **Abstract:** An attack detection device (120) comprises an abnormal pattern detection unit (123) that detects whether an abnormal pattern is included in a time series recognition score, said time series recognition score being time series data that was generated using a plurality of recognition scores that were respectively calculated using a plurality of items of

[続葉有]

MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告(条約第21条(3))

image data and indicate results in which an object was detected in each of the plurality of items of image data, said plurality of items of image data resulting from imaging a range within an imaging range at different times within an imaging time range, and said abnormal pattern occurring when an adversarial sample patch attack has been implemented against at least one of the plurality of items of image data.

(57) 要約:攻撃検知装置(１２０)は、撮影時間範囲内の互いに異なる時刻に撮影範囲内の範囲を撮影した複数の画像データそれぞれを用いて計算された複数の認識スコアであって、複数の画像データそれぞれにおいて物体を検知した結果を示す複数の認識スコアを用いて生成された時系列データである時系列認識スコアに、敵対的サンプルパッチ攻撃が複数の画像データの少なくともいずれかに対して実施された場合において生じる異常パターンが含まれているか否かを検知する異常パターン検知部(１２３)を備える。

明　細　書

発明の名称 ：

攻撃検知装置、敵対的サンプルパッチ検知システム、攻撃検知方法、及び、攻撃検知プログラム

技術分野

[0001] 本開示は、攻撃検知装置、敵対的サンプルパッチ検知システム、攻撃検知方法、及び、攻撃検知プログラムに関する。

背景技術

[0002] 入力画像における各オブジェクトの位置をバウンディングボックスで示し、各オブジェクトの種類をラベルとして示す物体検知のタスクにおいて、ニューラルネットワークを用いた深層学習の手法が近年、非常に高い精度を達成している。非特許文献１は、電子的な摂動が加えられた画像を印刷した敵対的サンプルパッチを物理的に配置し、配置された敵対的サンプルパッチを撮影した画像を入力した際に物体検知による検知を逃れる敵対的サンプルパッチ攻撃の手法を開示している。

先行技術文献

非特許文献

[0003] 非特許文献1：S. Thys et al., “Fooling automated surveillance cameras: adversarial patches to attack person detection”, CVPRW (Conference on Computer Vision and Pattern Recognition Workshops), IEEE (Institute of Electrical and Electronics Engineers) ／CVF (Computer Vision Foundation), 2019

発明の概要

発明が解決しようとする課題

[0004] 既存技術によれば、敵対的サンプルパッチ攻撃が行われた場合に、当該攻撃を検知することが困難であるという課題がある。

[0005] 本開示は、敵対的サンプルパッチ攻撃が行われた場合に、当該攻撃を検知することを目的とする。

課題を解決するための手段

[0006] 本開示に係る攻撃検知装置は、

撮影時間範囲内の互いに異なる時刻に撮影範囲内の範囲を撮影した複数の画像データそれぞれを用いて計算された複数の認識スコアであって、前記複数の画像データそれぞれにおいて物体を検知した結果を示す複数の認識スコアを用いて生成された時系列データである時系列認識スコアに、敵対的サンプルパッチ攻撃が前記複数の画像データの少なくともいずれかに対して実施された場合において生じる異常パターンが含まれているか否かを検知する異常パターン検知部

を備える。

発明の効果

[0007] 本開示に係る攻撃検知装置は、物体を検知した結果を示す認識スコアから成る時系列認識スコアに、敵対的サンプルパッチ攻撃が実施された場合において生じる異常パターンが含まれているか否かを検知する異常パターン検知部を備える。従って、本開示によれば、敵対的サンプルパッチ攻撃が行われた場合に、当該攻撃を検知することができる。

図面の簡単な説明

[0008] [図1]実施の形態１に係る敵対的サンプルパッチ検知システム１００のシステム構成例を示す図。

[図2]実施の形態１に係る物体検知装置１１０の機能構成例を示す図。

[図3]実施の形態１に係る攻撃検知装置１２０の機能構成例を示す図。

[図4]実施の形態１に係る物体検知装置１１０及び攻撃検知装置１２０の各々

のハードウェア構成例を示す図。

[図5]実施の形態１に係る敵対的サンプルパッチ検知システム１００の動作を示すフローチャート。

[図6]実施の形態１の変形例に係る物体検知装置１１０及び攻撃検知装置１２０の各々のハードウェア構成例を示す図。

発明を実施するための形態

[0009] 実施の形態の説明及び図面において、同じ要素及び対応する要素には同じ符号を付している。同じ符号が付された要素の説明は、適宜に省略又は簡略化する。図中の矢印はデータの流れ又は処理の流れを主に示している。また、「部」を、「回路」、「工程」、「手順」、「処理」又は「サーキットリー」に適宜読み替えてもよい。

[0010] 実施の形態１．

以下、本実施の形態について、図面を参照しながら詳細に説明する。

[0011] ＊＊＊構成の説明＊＊＊

図１は、本実施の形態に係る敵対的サンプルパッチ検知システム１００のシステム構成例を示している。敵対的サンプルパッチ検知システム１００は、物体検知装置１１０と攻撃検知装置１２０とを有する。物体検知装置１１０と攻撃検知装置１２０とは一体的に構成されてもよい。

物体検知装置１１０は、入力画像データｘを入力として受け取り、物体検知結果である認識スコアｙを出力する。

攻撃検知装置１２０は、認識スコアｙを入力として受け取り、敵対的サンプルパッチ攻撃の検知結果ｒとして、敵対的サンプルパッチ攻撃による認識スコアの異常パターンが検知された場合には敵対的サンプルパッチ攻撃を検知したことを示す結果、それ以外の場合には敵対的サンプルパッチ攻撃を検知していないことを示す結果を出力する。敵対的サンプルパッチ攻撃は、敵対的サンプル攻撃の一種であり、特にニューラルネットワーク等を用いた物体検知を逃れる攻撃である。異常パターンは、具体例として、認識スコアの値が、一定時間継続して物体検知の閾値を多少下回る値であるパターンであ

る。即ち、異常パターンは、時系列認識スコアが示す認識スコアの値が物体検知閾値未満異常検知閾値以上である時間が異常検知時間以上継続することを示すパターンである。現状の敵対的サンプルパッチを用いた攻撃では、認識スコアの値を、物体検知の閾値を下回るよう抑えることができるが、完全に０にすることはできない。そのため、この異常パターンは当該攻撃の検知において有効である。ただし、異常パターンはこのパターンのみに限定されない。

[0012] 図２は、物体検知装置１１０の機能構成例を示している。物体検知装置１１０は、データ入力部１１１と物体検知器１１２とデータ出力部１１３とを有する。

データ入力部１１１は、物体検知の対象である入力画像データｘを受け取り、受け取った入力画像データｘを物体検知器１１２に入力する。

物体検知器１１２は、入力された入力画像データｘを用いて認識スコアｙを算出し、算出した認識スコアｙをデータ出力部１１３に出力する。物体検知器１１２は、具体例としてニューラルネットワークによって構築された物体検知器１１２である。ニューラルネットワークは、具体例として、ＹＯＬＯ（Ｙｏｕ　Ｏｎｌｙ　Ｌｏｏｋ　Ｏｎｃｅ）、ＳＳＤ（Ｓｉｎｇｌｅ　Ｓｈｏｔ　Ｍｕｌｔｉｂｏｘ　Ｄｅｔｅｃｔｏｒ）、又はＦａｓｔｅｒ　Ｒ－ＣＮＮ（Ｒｅｇｉｏｎ－ｂａｓｅｄ　Ｃｏｎｖｏｌｕｔｉｏｎａｌ　Ｎｅｕｒａｌ　Ｎｅｔｗｏｒｋｓ）等である。物体検知器１１２は、入力された画像に映る各オブジェクトに対応するバウンディングボックスの位置を表す座標と、各バウンディングボックス内のオブジェクトの種類及び確信度を示す確率を認識スコアとして出力する。物体検知器１１２は、複数の画像データそれぞれを用いて複数の認識スコアを計算する。物体検知器１１２は物体検知部とも呼ばれる。

データ出力部１１３は、物体検知器１１２が計算した認識スコアｙを出力する。

[0013] 図３は、攻撃検知装置１２０の機能構成例を示している。攻撃検知装置１

２０は、データ入力部１２１と認識スコア集積部１２２と異常パターン検知部１２３とデータ出力部１２４とを有する。

データ入力部１２１は、認識スコアｙを受け取り、受け取った認識スコアｙを認識スコア集積部１２２に入力する。

認識スコア集積部１２２は、入力された認識スコアｙを時系列認識スコアＹ’に追加することによって時系列認識スコアＹを生成し、生成した時系列認識スコアＹを異常パターン検知部１２３に入力する。時系列認識スコアＹは更新された時系列認識スコアＹ’に当たる。時系列認識スコアＹ’は、新たな認識スコアｙが入力されるまでに入力された認識スコアｙを集積して生成した時系列データである。認識スコア集積部１２２は、認識スコア集積部１２２に認識スコアｙが入力される度に時系列認識スコアＹを異常パターン検知部１２３に入力しなくてもよく、一定時間毎に異常パターン検知部１２３が認識スコア集積部１２２から最新の時系列認識スコアＹを取り出し、取り出した時系列認識スコアＹを用いてもよい。時系列認識スコアは、複数の認識スコアを用いて生成された時系列データである。複数の認識スコアは、撮影時間範囲内の互いに異なる時刻に撮影範囲内の範囲を撮影した複数の画像データそれぞれを用いて計算され、また、当該複数の画像データそれぞれにおいて物体を検知した結果を示す。

異常パターン検知部１２３は、時系列認識スコアＹに対して事前に指定された異常パターンとのマッチングを行う。異常パターン検知部１２３は、検知結果ｒとして、時系列認識スコアＹが異常パターンに合致する場合には検知したことを示す結果、時系列認識スコアＹが異常パターンに合致しない場合には敵対的サンプルパッチ攻撃を検知していないことを示す結果をデータ出力部１２４に入力する。異常パターン検知部１２３は、通常時の認識スコアの推移と異なる推移である異常パターンを検知した際に敵対的サンプルパッチによる攻撃が行われたと判定する。異常パターン検知部１２３は、時系列認識スコアに、敵対的サンプルパッチ攻撃が複数の画像データの少なくともいずれかに対して実施された場合において生じる異常パターンが含まれて

いるか否かを検知する。

データ出力部１２４は、入力された検知結果ｒを出力する。

[0014] 図４は、本実施の形態に係る物体検知装置１１０及び攻撃検知装置１２０の各々のハードウェア資源の一例を示す図である。物体検知装置１１０及び攻撃検知装置１２０の各々は、コンピュータから成り、また、複数のコンピュータから成ってもよい。

[0015] 物体検知装置１１０及び攻撃検知装置１２０の各々は、プロセッサ１１を備えている。プロセッサ１１は、バス１２を介してＲＯＭ１３と、ＲＡＭ１４と、通信ボード１５と、表示装置であるディスプレイ５１と、キーボード５２と、マウス５３と、ドライブ５４と、磁気ディスク装置２０等のハードウェアデバイスと接続され、これらのハードウェアデバイスを制御する。プロセッサ１１は、演算処理を行うＩＣ（Ｉｎｔｅｇｒａｔｅｄ Ｃｉｒｃｕｉｔ）であり、具体例として、ＣＰＵ（Ｃｅｎｔｒａｌ Ｐｒｏｃｅｓｓｉｎｇ Ｕｎｉｔ）、ＤＳＰ（Ｄｉｇｉｔａｌ Ｓｉｇｎａｌ Ｐｒｏｃｅｓｓｏｒ）、又はＧＰＵ（Ｇｒａｐｈｉｃｓ Ｐｒｏｃｅｓｓｉｎｇ Ｕｎｉｔ）である。物体検知装置１１０及び攻撃検知装置１２０の各々は、複数のプロセッサを備えてもよい。複数のプロセッサは、プロセッサ１１の役割を分担する。

[0016] ドライブ５４は、ＦＤ（Ｆｌｅｘｉｂｌｅ Ｄｉｓｋ Ｄｒｉｖｅ）、ＣＤ（Ｃｏｍｐａｃｔ Ｄｉｓｃ）、又はＤＶＤ（Ｄｉｇｉｔａｌ Ｖｅｒｓａｔｉｌｅ Ｄｉｓｃ）等の記憶媒体を読み書きする装置である。

[0017] ＲＯＭ１３と、ＲＡＭ１４と、磁気ディスク装置２０と、ドライブ５４との各々は、記憶装置の一例である。記憶装置はコンピュータから独立していてもよい。

[0018] キーボード５２と、マウス５３と、通信ボード１５との各々は入力装置の一例である。ディスプレイ５１及び通信ボード１５は出力装置の一例である。

[0019] 通信ボード１５は、有線又は無線で、ＬＡＮ（Ｌｏｃａｌ Ａｒｅａ Ｎ

ｅｔｗｏｒｋ）、インターネット、又は電話回線等の通信網に接続している。通信ボード１５は、具体例として、通信チップ又はＮＩＣ（Ｎｅｔｗｏｒｋ　Ｉｎｔｅｒｆａｃｅ　Ｃａｒｄ）から成る。

[0020] 磁気ディスク装置２０は、ＯＳ（オペレーティングシステム）２１と、プログラム群２２と、ファイル群２３とを記憶している。

[0021] プログラム群２２は、本実施の形態において各部又は各器として説明する機能を実行するプログラムを含む。プログラムは、プロセッサ１１により読み出され実行される。即ち、プログラムは、各部又は各器としてコンピュータを機能させるものであり、また、各部又は各器の手順又は方法をコンピュータに実行させるものである。

本明細書に記載されているいずれのプログラムも、コンピュータが読み取り可能な不揮発性の記録媒体に記録されていてもよい。不揮発性の記録媒体は、具体例として、光ディスク又はフラッシュメモリである。本明細書に記載されているいずれのプログラムも、プログラムプロダクトとして提供されてもよい。

[0022] ファイル群２３は、本実施の形態において説明する各部又は各器で使用される各種データを含む。

[0023] ＊＊＊動作の説明＊＊＊

物体検知装置１１０の動作手順は、物体検知方法に相当する。また、物体検知装置１１０の動作を実現するプログラムは、物体検知プログラムに相当する。攻撃検知装置１２０の動作手順は、攻撃検知方法に相当する。また、攻撃検知装置１２０の動作を実現するプログラムは、攻撃検知プログラムに相当する。

[0024] 図５は、本実施の形態に係る敵対的サンプルパッチ検知システム１００の処理の一例を示すフローチャートである。本図を参照して敵対的サンプルパッチ検知システム１００の処理を説明する。

[0025] （ステップＳ１１）

データ入力部１１１は、入力画像データｘを受け取り、受け取った入力画

像データｘを物体検知器１１２に入力する。

[0026]（ステップＳ１２）

物体検知器１１２は、入力された入力画像データｘを用いて認識スコアｙを算出する。

[0027]（ステップＳ１３）

データ出力部１１３は、算出された認識スコアｙを出力する。

[0028]（ステップＳ１４）

データ入力部１２１は、認識スコアｙを受け取り、受け取った認識スコアｙを認識スコア集積部１２２に入力する。

[0029]（ステップＳ１５）

認識スコア集積部１２２は、入力された認識スコアｙを時系列認識スコアＹ’に追加することによって時系列認識スコアＹ’を時系列認識スコアＹに更新し、更新した時系列認識スコアＹを異常パターン検知部１２３に入力する。

[0030]（ステップＳ１６）

異常パターン検知部１２３は、入力された時系列認識スコアＹが事前に指定された異常パターンに合致するか否かを判定する。

異常パターンが時系列認識スコアＹに合致する場合、攻撃検知装置１２０はステップＳ１７に進む。それ以外の場合、攻撃検知装置１２０はステップＳ１８に進む。

[0031]（ステップＳ１７）

データ出力部１２４は、検知結果ｒとして、敵対的サンプルパッチ攻撃を検知したことを示す結果を出力する。

[0032]（ステップＳ１８）

データ出力部１２４は、検知結果ｒとして、敵対的サンプルパッチ攻撃を検知していないことを示す結果を出力する。

[0033]＊＊＊実施の形態１の効果の説明＊＊＊

以上のように、本実施の形態によれば、時系列認識スコアが異常パターン

に合致するか否かを判定することによって敵対的サンプルパッチ攻撃を検知することができる。

[0034] ＊＊＊他の構成＊＊＊

＜変形例１＞

図６は、本変形例に係る物体検知装置１１０及び攻撃検知装置１２０の各々のハードウェア構成例を示している。

物体検知装置１１０及び攻撃検知装置１２０の各々は、プロセッサ１１、プロセッサ１１とＲＯＭ１３、プロセッサ１１とＲＡＭ１４、あるいはプロセッサ１１とＲＯＭ１３とＲＡＭ１４とに代えて、処理回路１８を備える。

処理回路１８は、物体検知装置１１０及び攻撃検知装置１２０の各々が備える各部の少なくとも一部を実現するハードウェアである。

処理回路１８は、専用のハードウェアであってもよく、また、磁気ディスク装置２０に格納されるプログラムを実行するプロセッサであってもよい。

[0035] 処理回路１８が専用のハードウェアである場合、処理回路１８は、具体例として、単一回路、複合回路、プログラム化したプロセッサ、並列プログラム化したプロセッサ、ＡＳＩＣ（Ａｐｐｌｉｃａｔｉｏｎ Ｓｐｅｃｉｆｉｃ Ｉｎｔｅｇｒａｔｅｄ Ｃｉｒｃｕｉｔ）、ＦＰＧＡ（Ｆｉｅｌｄ Ｐｒｏｇｒａｍｍａｂｌｅ Ｇａｔｅ Ａｒｒａｙ）又はこれらの組み合わせである。

物体検知装置１１０及び攻撃検知装置１２０の各々は、処理回路１８を代替する複数の処理回路を備えてもよい。複数の処理回路は、処理回路１８の役割を分担する。

[0036] 物体検知装置１１０及び攻撃検知装置１２０の各々において、一部の機能が専用のハードウェアによって実現されて、残りの機能がソフトウェア又はファームウェアによって実現されてもよい。

[0037] 処理回路１８は、具体例として、ハードウェア、ソフトウェア、ファームウェア、又はこれらの組み合わせにより実現される。

プロセッサ１１とＲＯＭ１３とＲＡＭ１４と処理回路１８とを、総称して

「プロセッシングサーキットリー」という。つまり、物体検知装置１１０及び攻撃検知装置１２０の各々の各機能構成要素の機能は、プロセッシングサーキットリーにより実現される。

[0038] ＊＊＊他の実施の形態＊＊＊

実施の形態１について説明したが、本実施の形態のうち、複数の部分を組み合わせて実施しても構わない。あるいは、本実施の形態を部分的に実施しても構わない。その他、本実施の形態は、必要に応じて種々の変更がなされても構わず、全体としてあるいは部分的に、どのように組み合わせて実施されても構わない。各部又は各器として説明するものは、ファームウェア、ソフトウェア、ハードウェア又はこれらの組み合わせのいずれで実装されても構わない。

なお、前述した実施の形態は、本質的に好ましい例示であって、本開示と、その適用物と、用途の範囲とを制限することを意図するものではない。フローチャート等を用いて説明した手順は、適宜変更されてもよい。

符号の説明

[0039] １１　プロセッサ、１２　バス、１３　ＲＯＭ、１４　ＲＡＭ、１５　通信ボード、１８　処理回路、２０　磁気ディスク装置、２１　ＯＳ、２２　プログラム群、２３　ファイル群、５１　ディスプレイ、５２　キーボード、５３　マウス、５４　ドライブ、１００　敵対的サンプルパッチ検知システム、１１０　物体検知装置、１１１　データ入力部、１１２　物体検知器、１１３　データ出力部、１２０　攻撃検知装置、１２１　データ入力部、１２２　認識スコア集積部、１２３　異常パターン検知部、１２４　データ出力部。

請求の範囲

[請求項1] 撮影時間範囲内の互いに異なる時刻に撮影範囲内の範囲を撮影した複数の画像データそれぞれを用いて計算された複数の認識スコアであって、前記複数の画像データそれぞれにおいて物体を検知した結果を示す複数の認識スコアを用いて生成された時系列データである時系列認識スコアに、敵対的サンプルパッチ攻撃が前記複数の画像データの少なくともいずれかに対して実施された場合において生じる異常パターンが含まれているか否かを検知する異常パターン検知部
を備える攻撃検知装置。

[請求項2] 前記異常パターンは、前記時系列認識スコアが示す認識スコアの値が物体検知閾値未満異常検知閾値以上である時間が異常検知時間以上継続することを示すパターンである請求項１に記載の攻撃検知装置。

[請求項3] 請求項１又は２に記載の攻撃検知装置と、
前記複数の画像データそれぞれを用いて前記複数の認識スコアを計算する物体検知部を備える物体検知装置と
を備える敵対的サンプルパッチ検知システム。

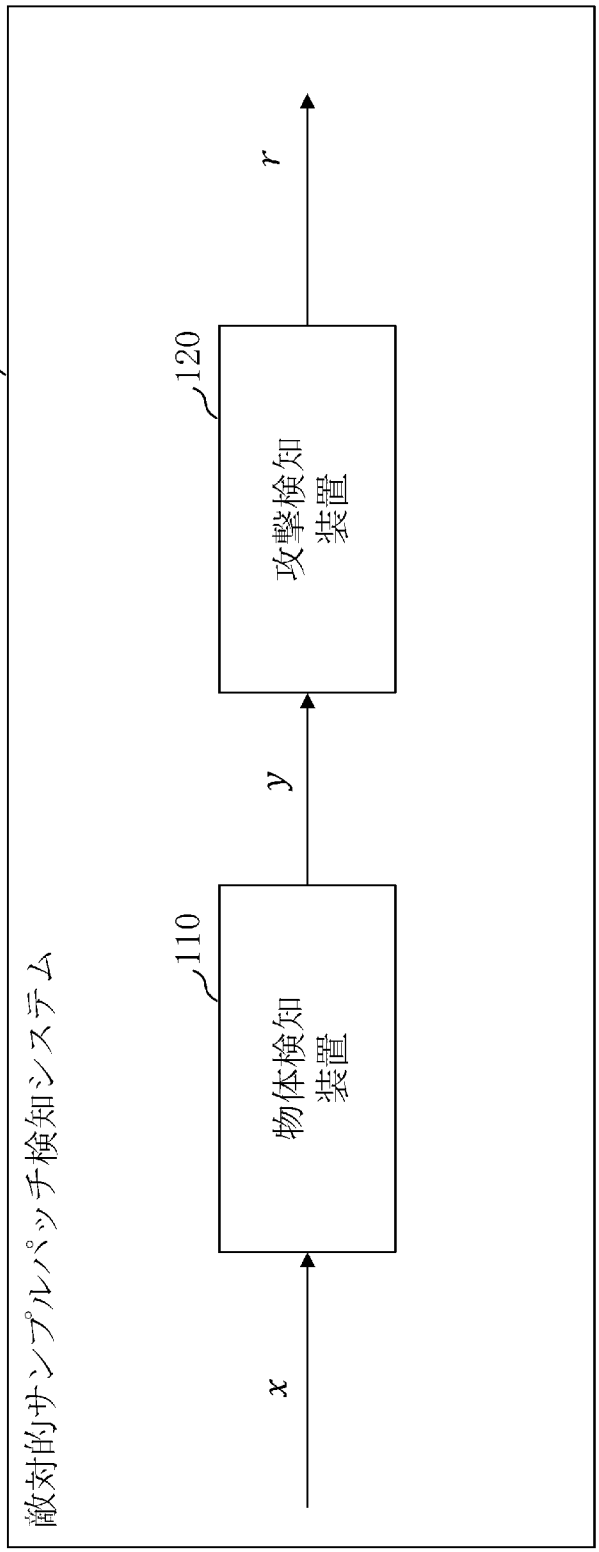
[請求項4] 撮影時間範囲内の互いに異なる時刻に撮影範囲内の範囲を撮影した複数の画像データそれぞれを用いて計算された複数の認識スコアであって、前記複数の画像データそれぞれにおいて物体を検知した結果を示す複数の認識スコアを用いて生成された時系列データである時系列認識スコアに、敵対的サンプルパッチ攻撃が前記複数の画像データの少なくともいずれかに対して実施された場合において生じる異常パターンが含まれているか否かを検知する攻撃検知方法。

[請求項5] 撮影時間範囲内の互いに異なる時刻に撮影範囲内の範囲を撮影した複数の画像データそれぞれを用いて計算された複数の認識スコアであって、前記複数の画像データそれぞれにおいて物体を検知した結果を示す複数の認識スコアを用いて生成された時系列データである時系列認識スコアに、敵対的サンプルパッチ攻撃が前記複数の画像データの

少なくともいずれかに対して実施された場合において生じる異常パターンが含まれているか否かを検知する異常パターン検知処理
をコンピュータである攻撃検知装置に実行させる攻撃検知プログラム。

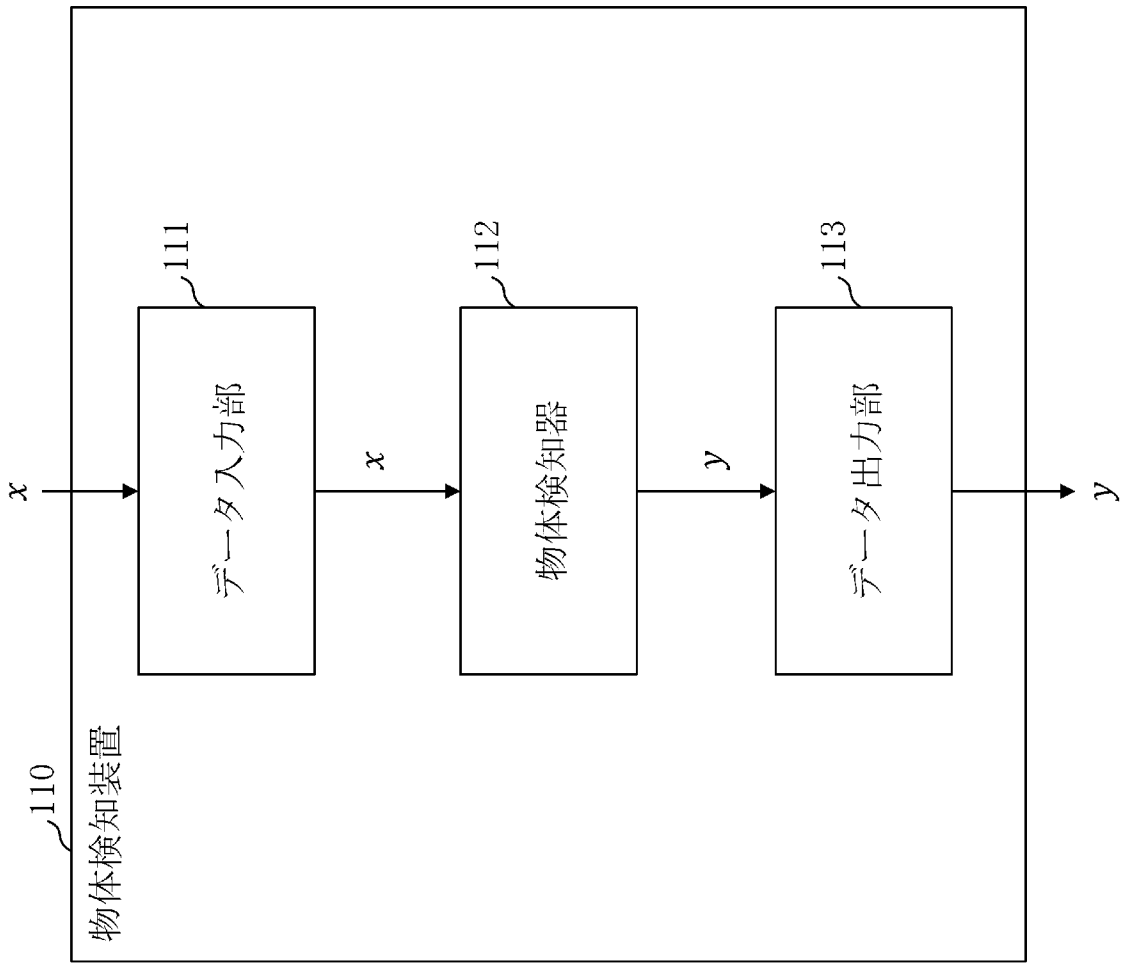
[図1]

図1



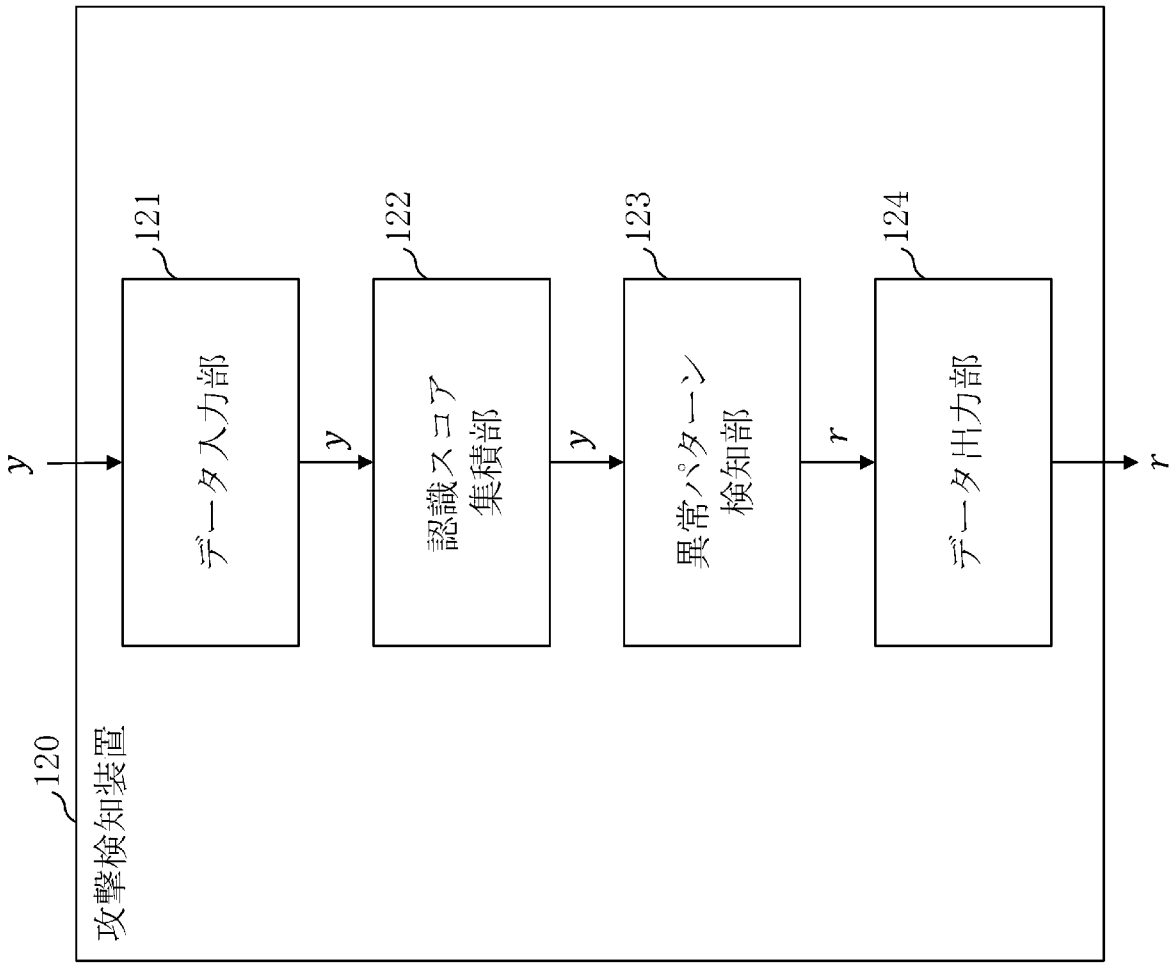
[図2]

図2



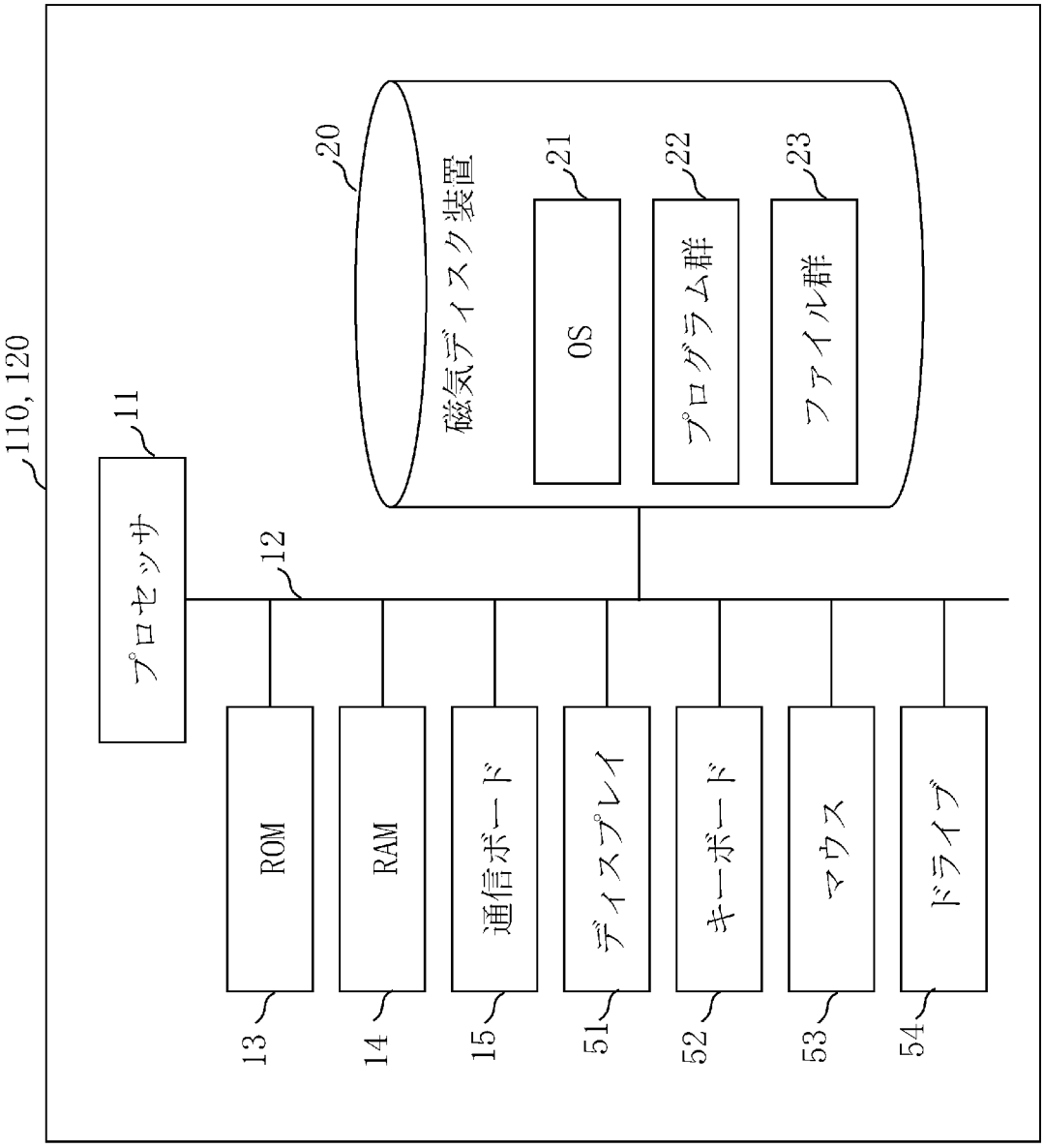
[図3]

図3


攻撃検知装置
120
データ入力部
121
認識スコア集積部
122
異常パターン検知部
123
データ出力部
124
y
y
y
r
r

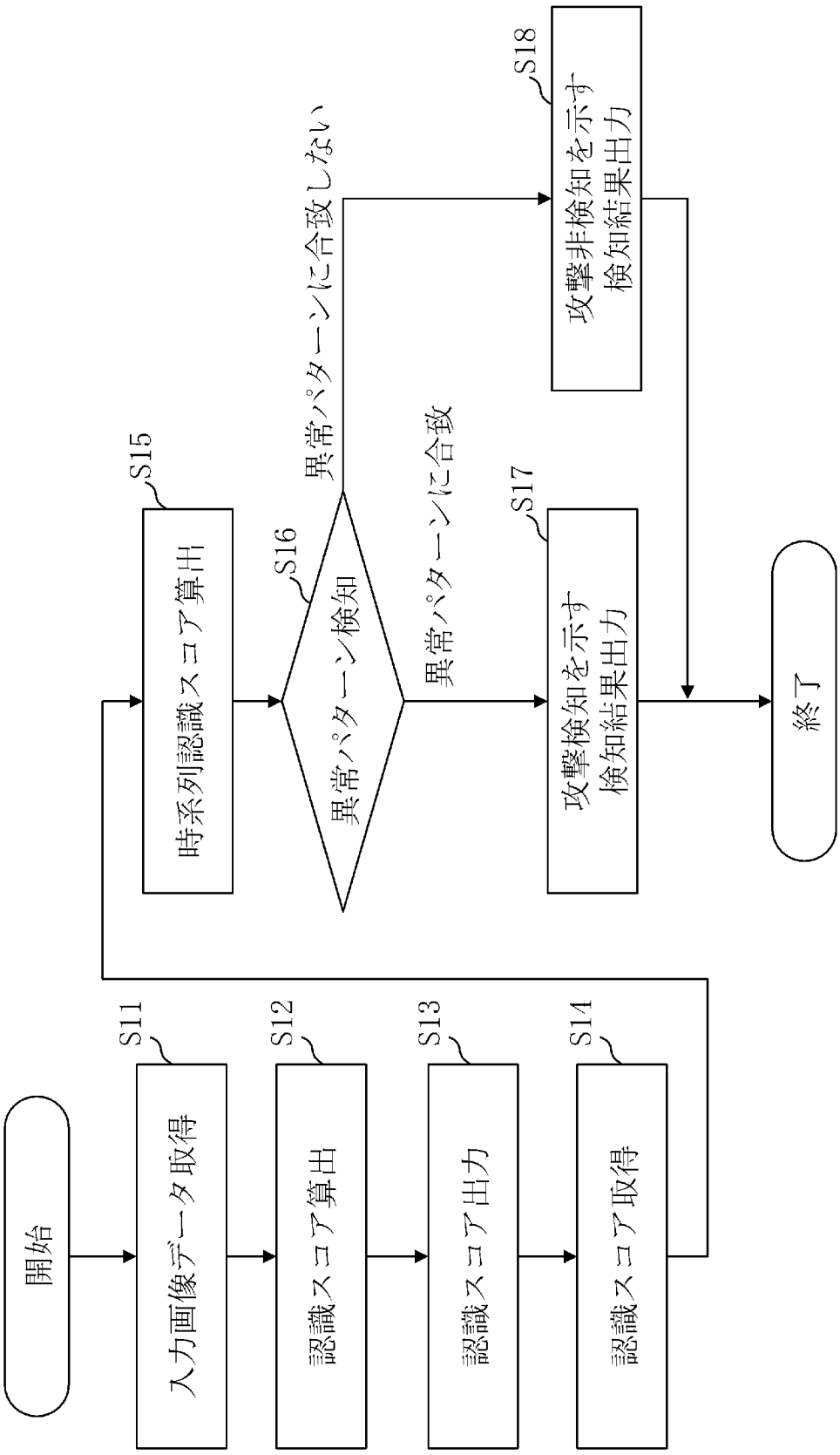
[図4]

図4


110, 120
11
プロセッサ
12
20
磁気ディスク装置
21
OS
22
プログラム群
23
ファイル群
ROM
RAM
通信ボード
ディスプレイ
キーボード
マウス
ドライブ
13
14
15
51
52
53
54

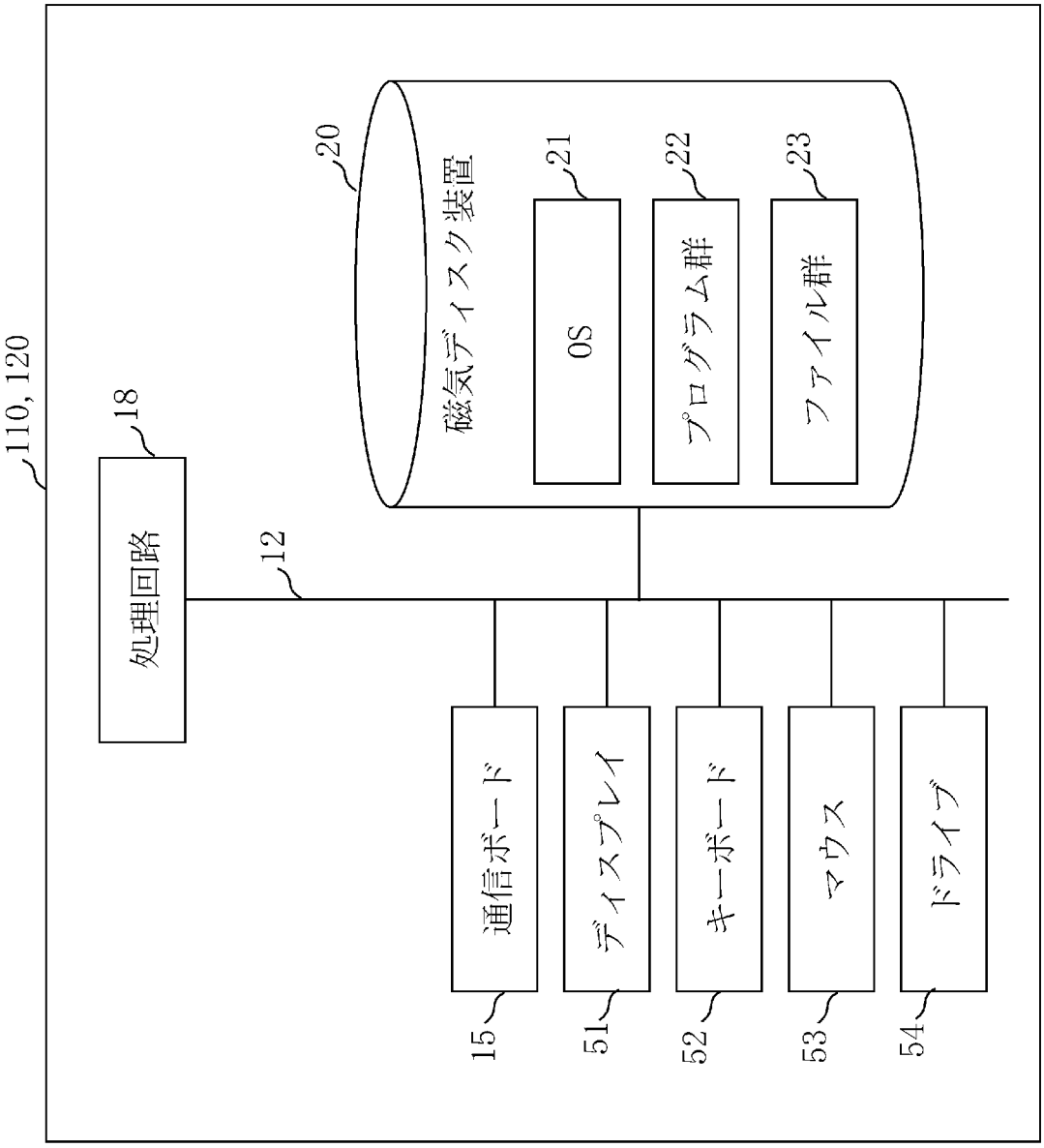
[図5]

図5


開始
S11
入力画像データ取得
S12
認識スコア算出
S13
認識スコア出力
S14
認識スコア取得
S15
時系列認識スコア算出
S16
異常パターン検知
異常パターンに合致しない
異常パターンに合致
S17
攻撃検知を示す
検知結果出力
S18
攻撃非検知を示す
検知結果出力
終了

[図6]

図6


110, 120
18
処理回路
12
20
磁気ディスク装置
21
OS
22
プログラム群
23
ファイル群
15
通信ボード
51
ディスプレイ
52
キーボード
53
マウス
54
ドライブ

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2021/022916

A. CLASSIFICATION OF SUBJECT MATTER
G06N 3/02(2006.01)i
FI: G06N3/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06N3/02

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2021
Registered utility model specifications of Japan	1996-2021
Published registered utility model applications of Japan	1994-2021

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	XIAO, C. et al. AdvIT: Adversarial Frames Identifier Based on Temporal Consistency in Videos. Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV)., 27 October 2019, pp. 3967-3976, Retrieved from the Internet: <URL:https://ieeexplore.ieee.org/document/9010733>, [Retrieved on 27 August 2021], DOI:10.1109/ICCV.2019.00407, in particular, sections 2-5	1, 3-5
A	entire text	2
A	DOSHI, K. et al. Continual Learning for Anomaly Detection in Surveillance Videos. arXiv.org [online]. arXiv:2004.07941., 15 April 2020, Retrieved from the Internet:<URL:https://arxiv.org/abs/2004.07941>, [Retrieved on 31 August 2021], entire text	2

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 August 2021 (27.08.2021)	07 September 2021 (07.09.2021)

Name and mailing address of the ISA/	Authorized officer
Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2021/022916

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JI, N. et al. Adversarial YOLO: Defense Human Detection Patch Attacks via Detecting Adversarial Patches. arXiv.org [online]. arXiv:2103.08860., 16 March 2021, Retrieved from the Internet:<URL:https://arxiv.org/abs/2103.08860>, [Retrieved on 31 August 2021], entire text	2
A	JP 2019-125867 A (PANASONIC IP MAN CORP) 25 July 2019 (2019-07-25) entire text, all drawings	2

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2021/022916

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
JP 2019-125867 A	25 Jul. 2019	US 2019/0217869 A1 whole document	

国際調査報告

国際出願番号
PCT/JP2021/022916

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

G06N 3/02(2006.01)i
FI: G06N3/02

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

G06N3/02

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０２１年
日本国実用新案登録公報	１９９６－２０２１年
日本国登録実用新案公報	１９９４－２０２１年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献のカテゴリー*	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
X	XIAO, C., et al., AdvIT: Adversarial Frames Identifier Based on Temporal Consistency in Videos, Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019.10.27, pp. 3967-3976, Retrieved from the Internet: <URL: https://ieeexplore.ieee.org/document/9010733> [Retrieved on 2021-08-27], DOI: 10.1109/ICCV.2019.00407 特に2-5節	1, 3-5
A	全文	2
A	DOSHI, K. et al., Continual Learning for Anomaly Detection in Surveillance Videos, arXiv.org [online], arXiv:2004.07941, 2020.04.15, Retrieved from the Internet: <URL: https://arxiv.org/abs/2004.07941> [Retrieved on 2021-08-31] 全文	2
A	JI, N., et al., Adversarial YOLO: Defense Human Detection Patch Attacks via Detecting Adversarial Patches, arXiv.org [online], arXiv:2103.08860, 2021.03.16, Retrieved from the Internet: <URL: https://arxiv.org/abs/2103.08860> [Retrieved on 2021-08-31] 全文	2

☑ C欄の続きにも文献が列挙されている。

☑ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー
"A" 特に関連のある文献ではなく、一般的技術水準を示すもの
"E" 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
"L" 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
"O" 口頭による開示、使用、展示等に言及する文献
"P" 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
"T" 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの
"X" 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
"Y" 特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
"&" 同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
27.08.2021	07.09.2021

名称及びあて先
日本国特許庁(ISA/JP)
〒100-8915
日本国
東京都千代田区霞が関三丁目４番３号

権限のある職員（特許庁審査官）
金木　陽一 5B 4876
電話番号 03-3581-1101 内線 3545

C.　関連すると認められる文献

引用文献のカテゴリー*	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
A	JP 2019-125867 A (パナソニックＩＰマネジメント株式会社) 25.07.2019 (2019-07-25) 全文，全図	2

国際調査報告
パテントファミリーに関する情報

国際出願番号
PCT/JP2021/022916

引用文献	公表日	パテントファミリー文献	公表日
JP 2019-125867 A	25.07.2019	US 2019/0217869 A1 Whole document	