



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년09월09일
(11) 등록번호 10-2154335
(24) 등록일자 2020년09월03일

(51) 국제특허분류(Int. Cl.)
G16H 50/20 (2018.01) A61B 5/145 (2006.01)
G01N 15/14 (2006.01) G06N 20/00 (2019.01)
G16H 50/70 (2018.01)
(52) CPC특허분류
G16H 50/20 (2018.01)
A61B 5/145 (2013.01)
(21) 출원번호 10-2019-0129523
(22) 출원일자 2019년10월18일
심사청구일자 2019년11월21일
(56) 선행기술조사문헌
JP2007132921 A
(뒷면에 계속)

(73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
주식회사 아이티메딕
경기도 용인시 기흥구 흥덕4로30번길 24 ,502호 (영덕동)
(72) 발명자
송재우
서울특별시 서초구 나루터로4길 28, 309동 104호 (잠원동, 신반포8차아파트)
이주범
경기도 수원시 영통구 범조로 38, 101동 203호(하동, 광고 더샵 레이크파크)
(74) 대리인
특허법인우인

전체 청구항 수 : 총 17 항

심사관 : 김미미

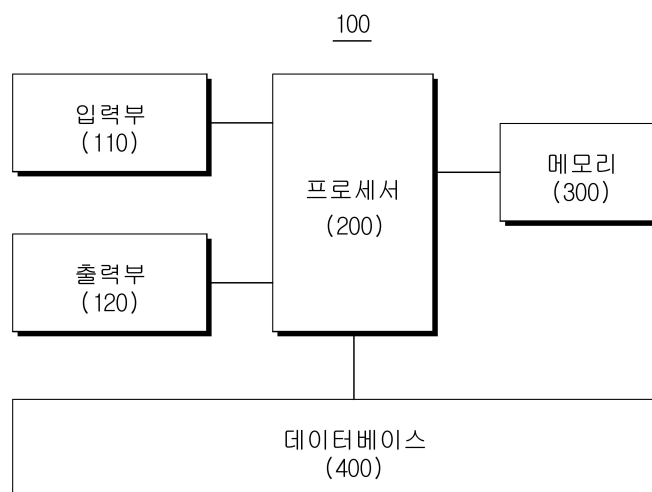
(54) 발명의 명칭 생체 추출 데이터를 전처리하여 질병을 판단하는 방법 및 그를 위한 장치

(57) 요약

생체 추출 데이터를 전처리하여 질병을 판단하는 방법 및 그를 위한 장치를 개시한다.

본 발명의 실시예에 따른 하나 이상의 프로세서 및 상기 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 컴퓨팅 디바이스에 의해 수행되는 질병 판단 방법에 있어서, 상기 컴퓨팅 디바이스는, 진단 대상의 혈액에서 추출된 생체 추출 데이터를 획득하는 데이터 획득 단계; 상기 생체 추출 데이터에 포함된 복수 개의 파라미터를 기반으로 생성된 초기 데이터를 복수의 채널 각각에 대한 좌표값으로 변형하고, 변형된 데이터를 학습용 데이터로 재구성하는 데이터 전처리 단계; 상기 재구성된 데이터에서 특징값을 추출하고, 상기 특징값을 분류하여 학습을 수행하는 데이터 학습 단계; 및 학습된 상기 특징값을 이용하여 특정 질병을 판단하는 질병 판단 단계를 수행할 수 있다.

대표도 - 도2



(52) CPC특허분류

A61B 5/7264 (2013.01)

A61B 5/7275 (2013.01)

G01N 15/14 (2013.01)

G06N 20/00 (2019.01)

G16H 50/70 (2018.01)

(56) 선행기술조사문헌

JP5025371 B2*

KR1020010017092 A*

KR1020180051844 A

US20180003634 A1*

*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호 2017M3A9G8084616

부처명 과학기술정보통신부

과제관리(전문)기관명 연구재단

연구사업명 원천기술개발사업 > 바이오·의료기술개발사업

연구과제명 데이터 마이닝/머신 러닝을 통한 임상검사실의 통상 검사 및 유세포 데이터 분석과

임상 예측 알고리즘의 개발

기 여 율 1/1

과제수행기관명 한양대학교(연세대학교 위탁)

연구기간 2019.01.01 ~ 2019.12.31

명세서

청구범위

청구항 1

하나 이상의 프로세서 및 상기 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 컴퓨팅 디바이스에 의해 수행되는 질병 판단 방법에 있어서,

상기 컴퓨팅 디바이스는,

진단 대상의 혈액에서 추출된 생체 추출 데이터를 획득하는 데이터 획득 단계;

상기 생체 추출 데이터에 포함된 복수 개의 파라미터를 기반으로 생성된 초기 데이터를 복수의 채널 각각에 대한 좌표값으로 변형하고, 변형된 데이터를 학습용 데이터로 재구성하는 데이터 전처리 단계;

상기 재구성된 학습용 데이터에서 특징값을 추출하고, 상기 특징값을 분류하여 학습을 수행하는 데이터 학습 단계; 및

학습된 상기 특징값을 이용하여 특정 질병을 판단하는 질병 판단 단계를 수행하되,

상기 데이터 전처리 단계는,

상기 초기 데이터에 포함된 일부 파라미터의 측정값들을 병합하여 상기 복수의 채널 각각에 대한 상기 좌표값을 포함하는 데이터로 변형하고, 상기 변형된 데이터와 상기 변형된 데이터 각각에 대한 카운트값을 포함하는 데이터 테이블을 생성하여 상기 학습용 데이터로 재구성하는 것을 특징으로 하는 질병 판단 방법.

청구항 2

제1항에 있어서,

상기 데이터 획득 단계는,

상기 혈액의 유세포 분석(Flow Cytometry)에 대한 결과를 나타내는 상기 생체 추출 데이터를 획득하며,

상기 생체 추출 데이터는, 연동하는 데이터베이스에 기 저장된 데이터에서 자동으로 수집되며, 기 설정된 정형화된 형식 또는 FCS(Flow Cytometry Standard) 형식의 데이터인 것을 특징으로 하는 질병 판단 방법.

청구항 3

제1항에 있어서,

상기 데이터 전처리 단계는,

상기 생체 추출 데이터에 포함된 검사 항목 채널의 복수 개의 파라미터의 전체 또는 일부 파라미터의 측정값을 이용하여 초기 데이터를 생성하는 초기 데이터 생성 단계;

상기 초기 데이터에 포함된 상기 일부 파라미터의 측정값들을 병합하여 상기 검사 항목 채널 각각에 대한 좌표값을 포함하는 데이터로 변형하고, 상기 변형된 데이터와 상기 변형된 데이터 각각에 대한 카운트값을 포함하는 데이터 테이블을 생성하는 데이터 변형 단계; 및

상기 데이터 테이블에 포함된 상기 변형된 데이터를 이용하여 기계학습용 데이터 테이블로 재구성하는 데이터 재구성 단계

를 포함하는 것을 특징으로 하는 질병 판단 방법.

청구항 4

제3항에 있어서,

상기 초기 데이터 생성 단계는,

상기 복수 개의 파라미터의 전체 또는 2 개 이상의 파라미터 각각에 대한 상기 측정값을 이용하여 상기 초기 데이터를 생성하는 것을 특징으로 하는 질병 판단 방법.

청구항 5

제3항에 있어서,

상기 데이터 변형 단계는,

상기 일부 파라미터의 측정값을 순차적 또는 기 설정된 순서로 병합하여 생성된 상기 좌표값을 포함하는 상기 변형된 데이터로 변형하는 것을 특징으로 하는 질병 판단 방법.

청구항 6

제3항에 있어서,

상기 데이터 변형 단계는,

상기 변형된 데이터에 포함된 상기 좌표값과 동일한 좌표값이 존재하는 경우, 동일한 좌표값을 삭제하고, 상기 좌표값에 대한 상기 카운트값을 기 설정된 단위로 증가하여 상기 카운트값을 갱신하고, 상기 변형된 데이터와 갱신된 상기 카운트값을 포함하는 상기 데이터 테이블을 생성하는 것을 특징으로 하는 질병 판단 방법.

청구항 7

제3항에 있어서,

상기 데이터 변형 단계는,

상기 초기 데이터에 포함된 전체 또는 일부 파라미터의 측정값들을 기 선정된 일정한 값으로 나눈 몫으로 치환하고, 데이터 유실을 막기 위해 각 몫에 소정의 값을 더하는 방식으로 이미지 깊이를 변형한 데이터와 상기 변형된 데이터 각각에 대한 카운트값을 포함하는 데이터 테이블을 생성하는 것을 특징으로 하는 질병 판단 방법.

청구항 8

제3항에 있어서,

상기 데이터 재구성 단계는,

상기 변형된 데이터에 포함된 상기 좌표값을 1 차원 좌표값으로 구성하고,

상기 1 차원 좌표값을 만드는 과정에서 좌표값이 존재하지 않는 부분을 0 값으로 채우거나 또는 좌표값이 존재하는 부분만 표시하는 방법을 사용하여 구성된 2 차원 또는 3 차원의 기계학습용 데이터 테이블로 재구성하는 것을 특징으로 하는 질병 판단 방법.

청구항 9

제3항에 있어서,

상기 데이터 학습 단계는,

합성망 알고리즘을 이용하여 상기 기계학습용 데이터 테이블에 포함된 상기 재구성된 데이터 내에서 특징값을 추출하고, 상기 특정 질병 별로 상기 특징값을 분류하여 학습하는 것을 특징으로 하는 질병 판단 방법.

청구항 10

제3항에 있어서,

상기 질병 판단 단계는,

상기 진단 대상에 대한 신규 정보가 입력된 경우, 상기 신규 정보와 상기 특정 질병에 대한 상기 특징값을 비교하여 질병을 판단하는 것을 특징으로 하는 질병 판단 방법.

청구항 11

생체 추출 데이터를 전처리하여 질병을 판단하는 장치로서,

하나 이상의 프로세서; 및

상기 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하며, 상기 프로그램들은 하나 이상의 프로세서에 의해 실행될 때, 상기 하나 이상의 프로세서들에서,

진단 대상의 혈액에서 추출된 생체 추출 데이터를 획득하는 데이터 획득 단계;

상기 생체 추출 데이터에 포함된 복수 개의 파라미터를 기반으로 생성된 초기 데이터를 복수의 채널 각각에 대한 좌표값으로 변형하고, 변형된 데이터를 학습용 데이터로 재구성하는 데이터 전처리 단계;

상기 재구성된 학습용 데이터에서 특징값을 추출하고, 상기 특징값을 분류하여 학습을 수행하는 데이터 학습 단계; 및

학습된 상기 특징값을 이용하여 특정 질병을 판단하는 질병 판단 단계를 포함하는 동작들을 수행하게 하되,

상기 데이터 전처리 단계는,

상기 초기 데이터에 포함된 일부 파라미터의 측정값들을 병합하여 상기 복수의 채널 각각에 대한 상기 좌표값을 포함하는 데이터로 변형하고, 상기 변형된 데이터와 상기 변형된 데이터 각각에 대한 카운트값을 포함하는 데이터 테이블을 생성하여 상기 학습용 데이터로 재구성하는 것을 특징으로 하는 질병 판단 장치.

청구항 12

제11항에 있어서,

상기 데이터 전처리 단계는,

상기 생체 추출 데이터에 포함된 검사 항목 채널의 복수 개의 파라미터의 전체 또는 일부 파라미터의 측정값을 이용하여 초기 데이터를 생성하는 초기 데이터 생성 단계;

상기 초기 데이터에 포함된 전체 또는 일부 파라미터의 측정값들을 기 선정된 일정한 값으로 나눈 몫으로 치환하고, 데이터 유실을 막기 위해 각 몫에 소정의 값을 더하는 방식으로 이미지 깊이를 변형한 데이터와 상기 변형된 데이터 각각에 대한 카운트값을 포함하는 데이터 테이블을 생성하는 데이터 변형 단계; 및

상기 데이터 테이블에 포함된 상기 변형된 데이터를 이용하여 기계학습용 데이터 테이블로 재구성하는 데이터 재구성 단계

를 포함하는 것을 특징으로 하는 질병 판단 장치.

청구항 13

제12항에 있어서,

상기 초기 데이터 생성 단계는,

상기 복수 개의 파라미터의 전체 또는 적어도 2 개 이상의 파라미터 각각에 대한 상기 측정값을 이용하여 상기 초기 데이터를 생성하며,

상기 변형된 데이터에 포함된 상기 좌표값과 동일한 좌표값이 존재하는 경우, 동일한 좌표값을 삭제하고, 상기 좌표값에 대한 상기 카운트값을 기 설정된 단위로 증가하여 상기 카운트값을 갱신하고, 상기 변형된 데이터와 갱신된 상기 카운트값을 포함하는 상기 데이터 테이블을 생성하는 것을 특징으로 하는 질병 판단 장치.

청구항 14

제12항에 있어서,

상기 데이터 재구성 단계는,

상기 변형된 데이터에 포함된 상기 좌표값을 1 차원 좌표값으로 구성하고, 상기 1 차원 좌표값으로 구성하는 과정에서 좌표값이 존재하지 않는 부분을 0 값으로 채우거나 또는 좌표값이 존재하는 부분만 표시하는 방법을 사용하여 구성된 2 차원 또는 3 차원의 기계학습용 데이터 테이블로 재구성하는 것을 특징으로 하는 질병 판단 장

치.

청구항 15

제12항에 있어서,

상기 데이터 학습 단계는,

합성망 알고리즘을 이용하여 상기 기계학습용 데이터 테이블에 포함된 상기 재구성된 데이터 내에서 특징값을 추출하고, 상기 특정 질병 별로 상기 특징값을 분류하여 학습하는 것을 특징으로 하는 질병 판단 장치.

청구항 16

제11항에 있어서,

상기 질병 판단 단계는,

상기 진단 대상에 대한 신규 정보가 입력된 경우, 상기 신규 정보와 상기 특정 질병에 대한 상기 특징값을 비교하여 질병을 판단하는 것을 특징으로 하는 질병 판단 장치.

청구항 17

컴퓨터에 제1항 내지 제10항 중 어느 한 항에 따른 질병 판단 방법을 실행시키기 위하여 기록매체에 저장된 컴퓨터프로그램.

발명의 설명

기술 분야

[0001] 본 발명은 생체 추출 데이터를 전처리하여 질병을 판단하는 방법 및 그를 위한 장치에 관한 것이다.

배경 기술

[0002] 이 부분에 기술된 내용은 단순히 본 발명의 실시예에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.

[0003] CBC(Complete Blood Count)를 통한 임상병리 검사 데이터의 중의 하나인 FCS(Flow Cytometry Standard) 데이터는 유체역학적 기술이나 이미지 분석을 통해 분석 시료 내의 세포(또는 이와 유사한 물리적, 유체역학적, 광학적 성질을 갖는 입자, 이하, 세포로 기재)들을 하나씩 분리하고, 이를 광학적 방식 또는 이와 유사한 방식으로 탐지 하여 세포의 개수 및 각 세포의 광학적/생물학적 성질을 정량화하고 분류하는 유세포분석, 이미지세포분석 및 이러한 일반적 분석법을 응용한 세포분석 결과를 콘텐츠(Content)로 하는 일종의 생체지표 데이터를 의미한다. 이러한 데이터는 각종 질병군과 연관성을 찾기 위한 좋은 지표로 활용할 수 있다.

[0004] 일반적으로 FCS 데이터 분석하는 방법은 도 1에 도시된 바와 같이, 분석 대상이 되는 세포(군집)을 분석자의 학술적 지식을 기반으로 세밀하게 선택/분리하고, 선택한 세포들을 계수 하거나 측정된 광학적 성질(예: 광분산 강도, 형광 발광) 및 이와 연관된 생물학적 성질(예: 크기, 구조, 항원 표현형)을 추출해 내는 과정으로 구성된다.

[0005] 자동화된 분석 방식으로 이러한 분석과정을 재현하거나 가공되지 않은 FCS 데이터로부터 군집분석 등을 통해 직접 분석 대상 군집을 분리해내는 알고리즘이 연구되고 있다. 이러한, 분석 대상 군집을 분리하는 기술은 한국등록특허 제10-1926632호에 기재되어 있다. 하지만, 대량의 임상병리 검사 데이터를 기반으로 질병 및 임상 상황/징후와 연관되어있는 FCS 데이터의 패턴(Pattern)을 추출하기 어려우며, 각 FCS 데이터의 전반적/형태적 특성을 기반으로 각 분석 대상 시료에 내포된 생물학적/임상적 의미를 찾기 위해 기계학습을 적용하기는 쉽지 않다. 따라서, 형상화된 이미지 형태로 FCS 데이터를 변환하여 FCS 데이터를 질환 및 환자별로 체계적으로 분류하여 지속적으로 다양한 질병 및 환자의 개인화된 특성과의 연관성을 찾는 기술이 필요하다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 진단 대상들의 질병 및 경과 관찰 과정에 생성되는 임상정보 중의 하나인 FCS(Flow Cytometry Standard) 데이터를 이미지 분석 기계학습이 가능하도록 FCS 데이터를 전환하고, 전환된 데이터를 기반으로 진행된 기계학습을 통해 다양한 질병과 관련된 지표화된 패턴을 찾고, 진단 대상의 특정 질병을 판단하는 생체 추출 데이터를 전처리하여 질병을 판단하는 방법 및 그를 위한 장치를 제공하는 데 주된 목적이 있다.

과제의 해결 수단

[0007] 본 발명의 일 측면에 의하면, 상기 목적을 달성하기 위한 하나 이상의 프로세서 및 상기 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 컴퓨팅 디바이스에 의해 수행되는 질병 판단 방법에 있어서, 상기 컴퓨팅 디바이스는, 진단 대상의 혈액에서 추출된 생체 추출 데이터를 획득하는 데이터 획득 단계; 상기 생체 추출 데이터에 포함된 복수 개의 파라미터를 기반으로 생성된 초기 데이터를 복수의 채널 각각에 대한 좌표값으로 변형하고, 변형된 데이터를 학습용 데이터로 재구성하는 데이터 전처리 단계; 상기 재구성된 학습용 데이터에서 특징값을 추출하고, 상기 특징값을 분류하여 학습을 수행하는 데이터 학습 단계; 및 학습된 상기 특징값을 이용하여 특정 질병을 판단하는 질병 판단 단계를 수행할 수 있다.

[0008] 또한, 본 발명의 다른 측면에 의하면, 상기 목적을 달성하기 위한 생체 추출 데이터를 전처리하여 질병을 판단하는 질병 진단 장치는 하나 이상의 프로세서; 및 상기 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하며, 상기 프로그램들은 하나 이상의 프로세서에 의해 실행될 때, 상기 하나 이상의 프로세서들에서, 진단 대상의 혈액에서 추출된 생체 추출 데이터를 획득하는 데이터 획득 단계; 상기 생체 추출 데이터에 포함된 복수 개의 파라미터를 기반으로 생성된 초기 데이터를 복수의 채널 각각에 대한 좌표값으로 변형하고, 변형된 데이터를 학습용 데이터로 재구성하는 데이터 전처리 단계; 상기 재구성된 학습용 데이터에서 특징값을 추출하고, 상기 특징값을 분류하여 학습을 수행하는 데이터 학습 단계; 및 학습된 상기 특징값을 이용하여 특정 질병을 판단하는 질병 판단 단계를 포함하는 동작들을 수행할 수 있다.

발명의 효과

[0009] 이상에서 설명한 바와 같이, 본 발명은 임상 예측을 위한 FCS 데이터 기계학습 모델 개발을 촉진함으로써 종래의 단편적 수치 비교에 기반한 질환 진단 방식에서 벗어나 자동혈액분석검사 및 유세포분석 결과의 상황적, 통합적 해석을 가능하게 하고 이를 통해 보다 정확한 질병진단과 임상 상황 파악을 할 수 있는 효과가 있다.

[0010] 또한, 본 발명은 임상적 유용성을 갖는 FCS 데이터 패턴을 발굴함에 따라, 의사가 인식하지 못한 환자의 이상을 조기에 발견하여 신속히 진단 또는 환자 파악에 이르도록 할 수 있는 효과가 있다.

[0011] 또한, 본 발명은 질병 특이 검사에 비해 저렴한 자동혈액분석검사 시행을 통해서 질병의 진행 및 환자 상태의 변화를 추적할 수 있도록 함으로써 의료 자원 분배의 효율성을 높이는 데 기여할 수 있는 효과가 있다.

[0012] 또한, 본 발명은 분석자의 수기 작업에 의존하는 기존의 유세포분석 검사 결과판독을 자동화한 새로운 알고리즘 개발을 촉진함으로써 생물학 및 의학 연구를 보다 원활하게 할 수 있는 효과가 있다.

[0013] 또한, 본 발명은 FCS 데이터 전환 방식(전처리 방식)을 통해 새로운 의료 기계학습 분야를 발전시킬 수 있는 효과가 있다.

도면의 간단한 설명

[0014] 도 1은 종래의 생체 추출 데이터의 분석 동작을 설명하기 위한 예시도이다.

도 2는 본 발명의 실시예에 따른 생체 추출 데이터 기반의 질병 진단 장치를 개략적으로 나타낸 블록 구성도이다.

도 3은 본 발명의 실시예에 따른 질병 진단 장치에 프로세서의 동작 구성을 개략적으로 나타낸 블록 구성도이다.

도 4는 본 발명의 실시예에 따른 생체 추출 데이터 기반의 질병 판단 방법을 설명하기 위한 순서도이다.

도 5는 본 발명의 실시예에 따른 환자정보 및 생체 추출 데이터를 이용하여 질병을 판단하는 동작을 설명하기 위한 예시도이다.

도 6은 본 발명의 실시예에 따른 신경 네트워크를 이용하여 질병을 판단하는 동작을 설명하기 위한 블록도이다.

도 7은 본 발명의 실시예에 따른 생체 추출 데이터 기반의 질병 진단 동작을 설명하기 위한 예시도이다.

도 8은 본 발명의 실시예에 따른 생체 추출 데이터를 기반으로 초기 데이터를 생성하는 동작을 설명하기 위한 예시도이다.

도 9a 및 도 9b는 본 발명의 실시예에 따른 복수의 채널 각각의 초기 데이터를 나타낸 예시도이다.

도 10a 및 도 10b는 본 발명의 실시예에 따른 생체 추출 데이터를 기반으로 기본적인 데이터를 변형하는 동작을 설명하기 위한 예시도이다.

도 11은 본 발명의 실시예에 따른 생체 추출 데이터를 기반으로 데이터를 재구성하는 동작을 예시하여 설명하기 위한 테이블에 대한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0015] 이하, 본 발명의 바람직한 실시예를 첨부된 도면들을 참조하여 상세히 설명한다. 본 발명을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다. 또한, 이하에서 본 발명의 바람직한 실시예를 설명할 것이나, 본 발명의 기술적 사상은 이에 한정하거나 제한되지 않고 당업자에 의해 변형되어 다양하게 실시될 수 있음은 물론이다. 이하에서는 도면들을 참조하여 본 발명에서 제안하는 생체 추출 데이터를 전처리하여 질병을 판단하는 방법 및 그를 위한 장치에 대해 자세하게 설명하기로 한다.
- [0016] 도 2는 본 발명의 실시예에 따른 생체 추출 데이터 기반의 질병 진단 장치를 개략적으로 나타낸 블록 구성도이다.
- [0017] 본 실시예에 따른 질병 진단 장치(100)는 입력부(110), 출력부(120), 프로세서(200), 메모리(300) 및 데이터 베이스(400)를 포함한다. 도 2의 질병 진단 장치(100)는 일 실시예에 따른 것으로서, 도 2에 도시된 모든 블록이 필수 구성요소는 아니며, 다른 실시예에서 질병 진단 장치(100)에 포함된 일부 블록이 추가, 변경 또는 삭제될 수 있다. 한편, 질병 진단 장치(100)에 포함된 각 구성요소들은 각각 별도의 소프트웨어 장치로 구현되거나, 소프트웨어가 결합된 별도의 하드웨어 장치로 구현될 수 있다.
- [0018] 질병 진단 장치(100)는 FCS(Flow Cytometry Standard) 데이터를 자동으로 학습용 데이터로 전처리하고, 전처리된 데이터를 기계학습과 인공지능 진단 모델의 데이터로 활용하고, 기계학습을 통해서 다양한 질병들의 특징값(Feature)를 찾아내고, 특징값들과 질병과의 연관성을 파악하여 예측 가능한 진단 모델을 생성하거나 특정 질병을 판단하는 동작을 수행한다.
- [0019] 입력부(110)는 질병 진단 장치(100)를 제어하기 위한 데이터를 입력하거나 획득하는 수단을 의미한다. 입력부(110)는 프로세서(200)와 연동하여 다양한 형태의 제어신호를 입력하거나, 외부 장치와 연동하여 직접 데이터를 획득하여 프로세서(200)로 전달할 수도 있다.
- [0020] 출력부(120)는 프로세서(200)와 연동하여 데이터의 전처리 결과, 학습결과, 진단 결과 등 다양한 정보를 표시할 수 있다. 출력부(120)는 질병 진단 장치(100)에 구비된 디스플레이(미도시)를 통해 다양한 정보를 표시하는 것이 바람직하나 반드시 이에 한정되는 것은 아니다.
- [0021] 프로세서(200)는 메모리(300)에 포함된 적어도 하나의 명령어 또는 프로그램을 실행시키는 기능을 수행한다.
- [0022] 본 실시예에 따른 프로세서(200)는 입력부(110) 또는 데이터 베이스(400)로부터 획득한 생체 추출 데이터를 기반으로 데이터 전처리를 수행하고, 전처리된 데이터를 기반으로 질병 진단을 위한 기계학습을 수행한다. 또한, 프로세서(200)는 기 학습된 학습결과를 기반으로 진단 대상의 질병을 판단할 수 있다. 본 실시예에 따른 프로세서(200)의 자세한 동작은 도 3에서 설명하도록 한다.
- [0023] 메모리(300)는 프로세서(200)에 의해 실행 가능한 적어도 하나의 명령어 또는 프로그램을 포함한다. 메모리(300)는 생체 추출 데이터를 기반으로 데이터를 전처리하는 동작을 위한 명령어 또는 프로그램을 포함할 수 있다. 또한, 메모리(300)는 전처리된 데이터를 기반으로 기계학습을 수행하는 동작을 위한 명령어 또는 프로그램을 포함할 수 있다. 또한, 메모리(300)는 학습 결과를 기반으로 진단 대상의 질병을 판단하는 동작을 위한 명령어 또는 프로그램을 포함할 수 있다.
- [0024] 데이터 베이스(400)는 데이터베이스 관리 프로그램(DBMS)을 이용하여 컴퓨터 시스템의 저장공간(하드디스크 또

는 메모리)에 구현된 일반적인 데이터구조를 의미하는 것으로, 데이터의 검색(추출), 삭제, 편집, 추가 등을 자유롭게 행할 수 있는 데이터 저장형태를 뜻하는 것으로, 오라클(Oracle), 인포믹스(Infomix), 사이베이스(Sybase), DB2와 같은 관계형 데이터베이스 관리 시스템(RDBMS)이나, 겔스톤(Gemston), 오리온(Orion), O2 등과 같은 객체 지향 데이터베이스 관리 시스템(OODBMS) 및 엑셀론(Excelon), 타미노(Tamino), 세카이주(Sekaiju) 등의 XML 전용 데이터베이스(XML Native Database)를 이용하여 본 발명의 일 실시예의 목적에 맞게 구현될 수 있고, 자신의 기능을 달성하기 위하여 적당한 필드(Field) 또는 엘리먼트들을 가지고 있다.

- [0025] 본 실시예에 따른 데이터베이스(400)는 생체 추출 데이터와 관련된 정보를 저장하고, 생체 추출 데이터 및 생체 추출 데이터와 관련된 정보를 제공할 수 있다. 데이터베이스(400)에 저장된 생체 추출 데이터는 혈액의 유세포 분석(Flow Cytometry)에 대한 결과를 나타내는 데이터일 수 있다. 생체 추출 데이터는 기 설정된 정형화된 형식 또는 FCS(Flow Cytometry Standard) 형식의 데이터인 것이 바람직하나 반드시 이에 한정되는 것은 아니다.
- [0026] 데이터베이스(140)는 질병 진단 장치(100) 내에 구현되는 것으로 기재하고 있으나 반드시 이에 한정되는 것은 아니며, 별도의 데이터 저장장치로 구현될 수도 있다.
- [0027] 도 3은 본 발명의 실시예에 따른 질병 진단 장치에 프로세서의 동작 구성을 개략적으로 나타낸 블록 구성도이다.
- [0028] 본 실시예에 따른 질병 진단 장치(100)에 포함된 프로세서(200)는 데이터 획득부(210), 데이터 전처리부(220), 데이터 학습부(230) 및 질병 진단부(240)를 포함한다. 도 3의 프로세서(200)는 일 실시예에 따른 것으로서, 도 3에 도시된 모든 블록이 필수 구성요소는 아니며, 다른 실시예에서 프로세서(200)에 포함된 일부 블록이 추가, 변경 또는 삭제될 수 있다. 한편, 프로세서(200)에 포함된 각 구성요소들은 각각 별도의 소프트웨어 장치로 구현되거나, 소프트웨어가 결합된 별도의 하드웨어 장치로 구현될 수 있다.
- [0029] 데이터 획득부(210)는 진단 대상의 혈액에서 추출된 생체 추출 데이터를 획득하는 동작을 수행한다. 여기서, 생체 추출 데이터는 혈액의 유세포 분석(Flow Cytometry)에 대한 결과를 나타내는 데이터일 수 있다. 생체 추출 데이터는 기 설정된 정형화된 형식 또는 FCS(Flow Cytometry Standard) 형식의 데이터인 것이 바람직하나 반드시 이에 한정되는 것은 아니다.
- [0030] 데이터 획득부(210)는 프로세서(200)와 연동하는 입력부(110) 또는 데이터 베이스(400)를 통해 생체 추출 데이터를 획득할 수 있다. 여기서, 데이터 획득부(210)는 프로세서(200)와 연동하는 데이터 베이스(400)로부터 생체 추출 데이터를 획득하는 경우, 기 설정된 주기마다 자동으로 생체 추출 데이터를 수집하거나 입력부(110)를 통해 입력된 데이터 요청신호를 데이터 베이스(400)로 전송하여 생체 추출 데이터를 수집할 수 있다.
- [0031] 데이터 전처리부(220)는 생체 추출 데이터에 포함된 복수 개의 파라미터를 기반으로 생성된 초기 데이터를 복수의 채널 각각에 대한 좌표값으로 변형하고, 변형된 데이터를 학습용 데이터로 재구성하는 동작을 수행한다. 본 실시예에 따른 데이터 전처리부(220)는 초기 데이터 생성부(222), 데이터 변형부(224) 및 데이터 재구성부(226)를 포함한다.
- [0032] 초기 데이터 생성부(222)는 생체 추출 데이터에 포함된 검사 항목 채널의 복수 개의 파라미터의 전체 또는 일부 파라미터의 측정값을 이용하여 초기 데이터를 생성한다.
- [0033] 초기 데이터 생성부(222)는 복수 개의 파라미터 중 적어도 둘 이상의 파라미터 각각에 대한 상기 측정값을 이용하여 상기 초기 데이터를 생성할 수 있다.
- [0034] 데이터 변형부(224)는 초기 데이터에 포함된 전체 또는 일부 파라미터의 측정값들을 가공없이 병합하여 상기 검사 항목 채널 각각에 대한 좌표값을 포함하는 데이터로 변형하고, 변형된 데이터와 상기 변형된 데이터 각각에 대한 카운트값을 포함하는 데이터 테이블을 생성한다.
- [0035] 또한 데이터 변형부(224)는 초기 데이터에 포함된 전체 또는 일부 파라미터의 측정값들을 선정된 일정한 값(예: 4, 8, 32 등의 특정 값)으로 나눈 몫으로 치환하는 방법으로 데이터를 변형(이미지 깊이 변환)하고, 이때 발생하는 데이터의 유실을 막기 위해 각 몫에 소정의 값(예: 10)을 더하는 방법을 취한다.
- [0036] 이렇게 변형된 데이터와 상기 변형된 데이터 각각에 대한 카운트 값을 포함하는 데이터 테이블을 생성한다.
- [0037] 데이터 변형부(224)는 일부 파라미터의 측정값을 순차적 또는 기 설정된 순서로 병합하여 생성된 좌표값을 포함하는 변형된 데이터로 변형한다.
- [0038] 또한, 데이터 변형부(224)는 변형된 데이터에 포함된 좌표값과 동일한 좌표값이 존재하는 경우, 동일한 좌표값

을 삭제하고, 좌표값에 대한 상기 카운트값을 기 설정된 단위로 증가하여 카운트값을 갱신하고, 변형된 데이터와 갱신된 상기 카운트값을 포함하는 상기 데이터 테이블을 생성한다.

- [0039] 데이터 재구성부(226)는 데이터 테이블에 포함된 변형된 데이터를 이용하여 기계학습용 데이터 테이블로 재구성하는 동작을 수행한다.
- [0040] 데이터 재구성부(226)는 변형된 데이터에 포함된 좌표값을 1 차원 좌표값으로 구성하고, 1 차원 좌표값으로 구성하는 과정에서 좌표값이 존재하지 않는 부분을 0 값으로 채우거나, 또는 좌표값이 존재하는 부분만 표시하는 방법을 사용하여 $\prod_{i=1}^m n_i$ (n_i 는 기 설정된 기준 크기값 이상의 자연수) 형태의 기계학습용 이미지(데이터 테이블)로 재구성할 수 있다. 여기서, 재구성된 기계학습용 이미지(데이터 테이블)는 2 차원 또는 3 차원의 형태일 수 있다.
- [0041] 본 실시예에 따른 데이터 전처리부(220)는 질병 진단 장치(100) 내에 포함된 것으로 기재하고 있으나 반드시 이에 한정되는 것은 아니며, 질병 진단 장치(100)와 별도의 장치로 구현될 수 있다. 예를 들어, 데이터 전처리부(220)는 생체 추출 데이터를 진단을 위한 기계학습용 데이터로 변환하는 데이터 전처리장치(미도시)와 같은 별도의 장치로 구현될 수 있으며, 데이터 전처리장치(미도시)는 다양한 형태로 학습을 수행하여 질병을 판단하는 장치와 연동할 수 있다.
- [0042] 데이터 학습부(230)는 재구성된 학습용 데이터에서 특징값을 추출하고, 추출된 특징값을 분류하여 질병 진단을 위한 학습을 수행한다. 본 실시예에 따른 데이터 학습부(230)는 특징 추출부(232) 및 특징 분류부(234)를 포함한다.
- [0043] 특징 추출부(232)는 합성망 알고리즘을 이용하여 기계학습용 데이터 테이블에 포함된 상기 재구성된 데이터 내에서 특징값을 추출한다.
- [0044] 특징 분류부(234)는 특정 질병 별로 특징값을 분류하여 학습을 수행한다.
- [0045] 질병 진단부(240)는 학습된 특징값을 이용하여 특정 질병을 판단하는 동작을 수행한다. 질병 진단부(240)는 진단 대상에 대한 신규 정보가 입력된 경우, 신규 정보와 특정 질병에 대한 특징값을 비교하여 질병을 판단한다.
- [0046] 도 4는 본 발명의 실시예에 따른 생체 추출 데이터 기반의 질병 판단 방법을 설명하기 위한 순서도이다.
- [0047] 질병 진단 장치(100)는 진단 대상의 혈액에서 추출된 생체 추출 데이터를 획득한다(S410). 여기서, 생체 추출 데이터는 혈액의 유세포 분석(Flow Cytometry)에 대한 결과를 나타내는 데이터일 수 있다. 생체 추출 데이터는 기 설정된 정형화된 형식 또는 FCS(Flow Cytometry Standard) 형식의 데이터인 것이 바람직하나 반드시 이에 한정되는 것은 아니다.
- [0048] 질병 진단 장치(100)는 생체 추출 데이터를 기반으로 초기 데이터를 생성한다(S420). 질병 진단 장치(100)는 생체 추출 데이터에 포함된 검사 항목 채널의 복수 개의 파라미터의 전체 또는 일부 파라미터의 측정값을 이용하여 초기 데이터를 생성한다.
- [0049] 질병 진단 장치(100)는 초기 데이터에 포함된 데이터를 변형하여 데이터 테이블을 생성한다(S430). 질병 진단 장치(100)는 초기 데이터에 포함된 일부 파라미터의 측정값들을 병합하여 상기 검사 항목 채널 각각에 대한 좌표값을 포함하는 데이터로 변형하고, 변형된 데이터와 상기 변형된 데이터 각각에 대한 카운트값을 포함하는 데이터 테이블을 생성한다.
- [0050] 질병 진단 장치(100)는 데이터 테이블에 포함된 변형된 데이터를 재구성하여 기계학습용 데이터 테이블을 생성한다(S440).
- [0051] 질병 진단 장치(100)는 데이터 테이블에 포함된 변형된 데이터에 포함된 좌표값을 1 차원 좌표값으로 구성하고, 1 차원 좌표값으로 구성하는 과정에서 좌표값이 존재하지 않는 부분을 0 값으로 채우거나, 또는 좌표값이 존재하는 부분만 표시하는 방법을 사용하여 $\prod_{i=1}^m n_i$ (n_i 는 기 설정된 기준 크기값 이상의 자연수) 형태의 기계학습용 이미지(데이터 테이블)로 재구성할 수 있다.
- [0052] 질병 진단 장치(100)는 합성망 알고리즘을 이용하여 기계학습용 데이터 테이블에 포함된 재구성된 데이터 내에서 특징값을 추출한다(S450).
- [0053] 질병 진단 장치(100)는 특징값을 기반으로 학습을 수행하여 특정 질병 별로 특징값을 분류한다(S460).

- [0054] 질병 진단 장치(100)는 학습된 특징값을 이용하여 특정 질병을 판단한다(S470). 질병 진단 장치(100)는 진단 대상에 대한 신규 정보가 입력된 경우, 신규 정보와 특정 질병에 대한 특징값을 비교하여 질병을 판단한다.
- [0055] 도 4에서는 각 단계를 순차적으로 실행하는 것으로 기재하고 있으나, 반드시 이에 한정되는 것은 아니다. 다시 말해, 도 4에 기재된 단계를 변경하여 실행하거나 하나 이상의 단계를 병렬적으로 실행하는 것으로 적용 가능할 것이므로, 도 4는 시계열적인 순서로 한정되는 것은 아니다.
- [0056] 도 4에 기재된 본 실시예에 따른 질병 판단 방법은 애플리케이션(또는 프로그램)으로 구현되고 단말장치(또는 컴퓨터)로 읽을 수 있는 기록매체에 기록될 수 있다. 본 실시예에 따른 질병 판단 방법을 구현하기 위한 애플리케이션(또는 프로그램)이 기록되고 단말장치(또는 컴퓨터)가 읽을 수 있는 기록매체는 컴퓨팅 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치 또는 매체를 포함한다.
- [0057] 도 5는 본 발명의 실시예에 따른 환자정보 및 생체 추출 데이터를 이용하여 질병을 판단하는 동작을 설명하기 위한 예시도이다.
- [0058] 질병 진단 장치(100)에서 데이터 전처리부(220)는 기계학습을 위한 데이터 전처리를 수행한다.
- [0059] 진단 대상을 구분할 수 있는 환자 정보를 식명화 처리하고, 식명화된 정보의 임상 검사 결과를 데이터 전처리부(220)로 입력한다.
- [0060] 데이터 전처리부(220)는 기 설정된 엑셀 형식 또는 FCS 형식의 생체 추출 데이터를 획득하고, 생체 추출 데이터에 포함된 복수의 파라미터의 측정값을 벡터 기반의 좌표값으로 표현하여 초기 데이터를 생성한다.
- [0061] 데이터 전처리부(220)는 초기 데이터에 포함된 복수 개의 파라미터의 좌표값들을 병합하여 하나의 좌표값으로 변형하고, 변형된 데이터와 각각의 병합된 좌표값의 카운팅을 통해 데이터 테이블(데이터 프레임)을 생성한다. 데이터 전처리부(220)는 데이터베이스에 저장된 데이터를 READ 또는 WRITE하여 데이터 테이블을 갱신 처리한다.
- [0062] 데이터 전처리부(220)는 데이터 테이블에 포함된 변형된 데이터를 재구성하여 변환한다. 데이터 전처리부(220)는 질병 진단 장치(100)는 데이터 테이블에 포함된 변형된 데이터에 포함된 좌표값을 1 차원 좌표값으로 구성하고, 1 차원 좌표값으로 구성하는 과정에서 좌표값이 존재하지 않는 부분을 0 값으로 채우거나, 또는 좌표값이 존재하는 부분만 표시하는 방법을 사용하여 $\prod_{i=1}^m n_i$ (n_i 는 기 설정된 기준 크기값 이상의 자연수) 형태의 기계학습용 이미지(데이터 테이블)로 재구성할 수 있다.
- [0063] 데이터 전처리부(220)는 변환된 기계학습용 데이터 또는 기계학습용 데이터 테이블을 데이터 학습부(230)로 전달하여 특정 질병을 판단하기 위한 학습이 수행되도록 한다.
- [0064] 도 6은 본 발명의 실시예에 따른 신경 네트워크를 이용하여 질병을 판단하는 동작을 설명하기 위한 블록도이다.
- [0065] 데이터 학습부(230)는 데이터 전처리부(220)에서 구성된 기계학습용 데이터를 입력 데이터로 사용하여 이미지 학습의 과정을 수행한다.
- [0066] 데이터 학습부(230)는 이미지 학습의 과정을 통하여 입력 데이터에서 특징값(Feature)을 검출하는 동작을 수행한다. 여기서, 데이터 학습부(230)는 복수의 컨볼루션 레이어(Convolutional Layer) 기반의 합성곱 알고리즘 및 기타 다른 향상된 기계학습 알고리즘을 이용하여 입력 데이터의 특징값을 검출할 수 있다.
- [0067] 데이터 학습부(230)는 검출된 특징값들을 기반으로 학습을 진행하여 특정 질병의 특징값들의 분류를 수행한다.
- [0068] 질병 진단부(240)는 데이터 학습부(230)의 학습결과를 기반으로 질병의 진단을 수행할 수 있다. 질병 진단부(240)는 진단 대상에 대한 신규 데이터 또는 기계학습 이전의 데이터가 입력된 경우, 해당 데이터에서 기 학습된 특정 질병(예: 혈액암 등) 환자군에서 추출한 특징값의 유무를 분석하고, 특징값의 유무에 따라 특정 질병을 판단할 수 있다.
- [0069] 도 7은 본 발명의 실시예에 따른 생체 추출 데이터 기반의 질병 진단 동작을 설명하기 위한 예시도이다.
- [0070] 본 실시예에 따른 질병 진단 장치(100)는 컴퓨터 내의 진단장치(700)로 구현될 수 있다. 컴퓨터 내의 진단장치(700)는 데이터 처리부(710), 특징값 생성부(720), 인공지능부(730) 및 진단부(740)를 포함하여 구성될 수 있다.
- [0071] 데이터 처리부(710)는 생체 추출 데이터에 포함된 복수 개의 파라미터를 기반으로 생성된 초기 데이터를 복수의 채널 각각에 대한 좌표값으로 변형하고, 변형된 데이터를 기계학습용 데이터로 재구성하는 동작을 수행한다. 여

기서, 데이터 처리부(710)는 데이터 전처리부(220)의 전체 또는 일부 기능을 포함하는 형태로 구현될 수 있다.

- [0072] 특징값 생성부(720)는 합성망 알고리즘 및 기타 다른 향상된 기계학습 알고리즘을 이용하여 기계학습용 데이터 테이블에 포함된 재구성된 데이터 내에서 추출된 특징값을 생성할 수 있다. 여기서, 특징값 생성부(720)는 데이터 학습부(230)의 일부 기능을 포함하는 형태로 구현될 수 있다.
- [0073] 인공지능부(730)는 추출된 특징값을 기반으로 학습을 수행하고, 학습결과에 따라 특정 질병 별로 특징값을 분류한다. 여기서, 인공지능부(730)는 데이터 학습부(230)의 일부 기능을 포함하는 형태로 구현될 수 있다.
- [0074] 진단부(740)는 학습된 특징값을 이용하여 특정 질병을 판단한다. 진단부(740)는 진단 대상에 대한 신규 정보가 입력된 경우, 신규 정보와 특정 질병에 대한 특징값을 비교하여 질병을 판단한다. 여기서, 진단부(740)는 질병 진단부(240)의 전체 또는 일부 기능을 포함하는 형태로 구현될 수 있다.
- [0075] 도 8은 본 발명의 실시예에 따른 생체 추출 데이터를 기반으로 초기 데이터를 생성하는 동작을 설명하기 위한 예시도이다.
- [0076] 도 8의 (a)를 참조하면, 진단 대상의 혈액에서 추출된 생체 추출 데이터는 복수 개의 파라미터를 포함하고, 복수 개의 파라미터 각각은 측정값을 포함한다. 예를 들어, 자동 혈구 분석기를 통해 추출된 생체 추출 데이터는 환자 및 검체 그리고 분석 장비의 분석 모듈 별로 두 개 내지 네 개의 파일로 구분되며, 각 파일은 분석 파라미터 별 측정값이 도 8의 (a)와 같이 나열되어 있는 테이블 형식으로 구현될 수 있다.
- [0077] 예를 들어, 생체 추출 데이터는 4 개의 분석 파라미터를 이용한 4 차원의 좌표들로 이루어진 점들의 집합일 수 있다. 하지만, 이미지 표현을 통해 이해를 돕기 위하여 생체 추출 데이터에 포함된 네 개의 파라미터 중 세 개의 파라미터를 선택하고, 선택된 파라미터를 이용하여 도 8의 (b)와 같이 3 차원적 좌표점들로 표현하였다. 여기서, 질병 진단 장치(100)는 선택된 파라미터를 통해 데이터 전처리를 위한 초기 데이터를 생성할 수 있다.
- [0078] 도 9a 및 도 9b는 본 발명의 실시예에 따른 복수의 채널 각각의 초기 데이터를 나타낸 예시도이다.
- [0079] 생체 추출 데이터를 기반으로 3 차원의 좌표점들은 도 9a 및 도 9b와 같은 플롯(plot)으로 그래프화될 수 있다. 이러한 좌표점들의 플롯 양상은 환자/검체 별로 유사하지만 미세한 차이를 보인다. 예를 들어, 자동혈액분석장비는 2 개 내지 4 개의 채널을 통한 개별적 분석을 동시에 시행하기 때문에 한 검체 당 2 개 내지 4 개의 FCS 데이터를 생성할 수 있다.
- [0080] 도 9a 및 도 9b는 10 명의 환자에서 수집한 자동혈구분석 FCS 데이터의 4 개 파라미터들(FCS, FCSW, SSC, SFL: 4 차원) 중 세 개의 파라미터를 3 차원 좌표에 열거해 놓은 것이다. 각 채널 별로 10 개의 FCS 데이터 플롯을 열거하여 육안 비교가 가능하도록 나타내었다.
- [0081] 도 9a의 (a)는 WDF channel(자동혈구분석기 백혈구 분석 channel 중 하나)에 대한 플롯들이고, 도 9a의 (b)는 WPC channel(자동혈구분석기 백혈구 분석 channel 중 하나)에 대한 플롯들을 나타낸다. 도 9b의 (a)는 WNR channel(자동혈구분석기 백혈구 분석 channel)에 대한 플롯들이고, 도 9b의 (b)는 PLT-F channel(자동혈구분석기 혈소판 분석 channel 중 하나)에 대한 플롯들을 나타낸다. 도 9a 및 도 9b에 도시된 각 플롯은 유사한 군집(Clustering) 양상을 보이거나 자세한 분포 양상에서 미세한 차이를 보인다.
- [0082] 도 10a 및 도 10b는 본 발명의 실시예에 따른 생체 추출 데이터를 기반으로 기본적인 데이터를 변형하는 동작을 설명하기 위한 예시도이다.
- [0083] 도 10a는 초기 데이터를 변형하는 동작을 설명하기 위한 데이터 테이블을 나타낸다.
- [0084] 질병 진단 장치(100)는 초기 데이터(FCS 데이터)의 파라미터 각각의 측정값들을 병합하여 각각의 검사 항목값이 하나의 좌표값이 되도록 변형한다.
- [0085] 또한 진단장치(100)는 초기 데이터에 포함된 전체 또는 일부 파라미터의 측정값들을 선정된 일정한 값(예: 4, 8, 32 등의 특정 값)으로 나눈 몫으로 치환하는 방법으로 데이터를 변형(이미지 값이 변환)하고, 이때 발생하는 데이터의 유실을 막기 위해 각 몫에 소정의 값(예: 10)을 더하는 방법을 취한다.
- [0086] 또한, 질병 진단 장치(100)는 변형된 데이터와 상기 변형된 데이터 각각에 대한 카운트값을 포함하는 데이터 테이블을 생성한다.
- [0087] 질병 진단 장치(100)는 변형된 데이터에 포함된 좌표값과 동일한 좌표값이 존재하는 경우, 동일한 좌표값을 삭제하고 좌표값에 대한 카운트값을 기 설정된 단위로 증가하여 카운트값을 갱신하고, 변형된 데이터와 갱신된 카

운트값을 포함하는 데이터 테이블을 생성한다. 예를 들어, 질병 진단 장치(100)는 변형된 데이터의 좌표값이 1개일 경우 카운트값을 1로 부여하고, 동일한 좌표값이 존재하는 경우 해당 좌표값의 카운트값을 2로 부여하는 방식으로 새로운 데이터 테이블을 생성할 수 있다.

[0088] 질병 진단 장치(100)는 데이터 테이블을 통해, 좌표 공간 내 각 픽셀(pixel)에 해당하는 좌표점의 수를 계산할 수 있다. 도 10b의 (a)의 예시는 변형된 데이터에 포함된 좌표값을 그래프로 나타낸 것이고, 도 10b의 (b)는 데이터 테이블을 통해 좌표 공간 내 각 픽셀(pixel)에 해당하는 좌표점을 카운팅하는 동작을 나타낸다.

[0089] 도 11은 본 발명의 실시예에 따른 생체 추출 데이터를 기반으로 데이터를 재구성하는 동작을 예시하여 설명하기 위한 테이블에 대한 도면이다.

[0090] 질병 진단 장치(100)는 데이터 테이블에 포함된 좌표값과 카운트값을 병합하고, 병합된 값으로 1차원 좌표값을 구성한다. 이후, 질병 진단 장치(100)는 1차원 좌표값을 기반으로 데이터를 재구성하여 기계학습용 데이터 테이블을 생성한다. 예를들어, 질병 진단 장치(100)는 1차원 좌표값을 2차원 배열로 재구성할 수 있다.

[0091] 질병 진단 장치(100)는 변형된 데이터에 포함된 좌표값을 1차원 좌표값으로 구성하고, 1차원 좌표값으로 구성하는 과정에서 좌표값이 존재하지 않는 부분을 0값으로 채우거나, 또는 좌표값이 존재하는 부분만 표시하는 방법을 사용하여 $\prod_{i=1}^m n_i$ (n_i 는 기 설정된 기준 크기값 이상의 자연수) 형태의 기계학습용 데이터 테이블로 재구성할 수 있다. 예를 들어, 질병 진단 장치(100)는 도 11에 도시된 바와 같이, 12 * 12 크기의 기계학습용 데이터 테이블과 같이 데이터를 재구성할 수 있다. 여기서, 하나의 행은 하나의 좌표값 및 카운트값을 의미한다.

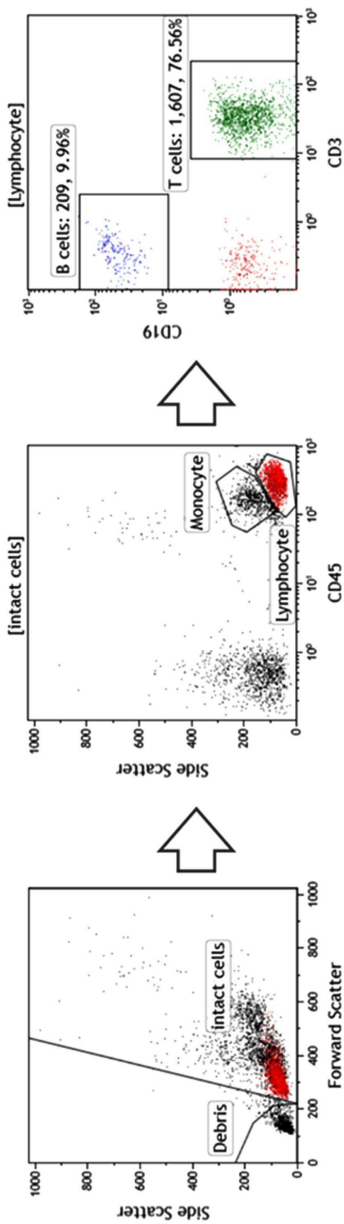
[0092] 이상의 설명은 본 발명의 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명의 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 발명의 실시예들은 본 발명의 실시예의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 발명의 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

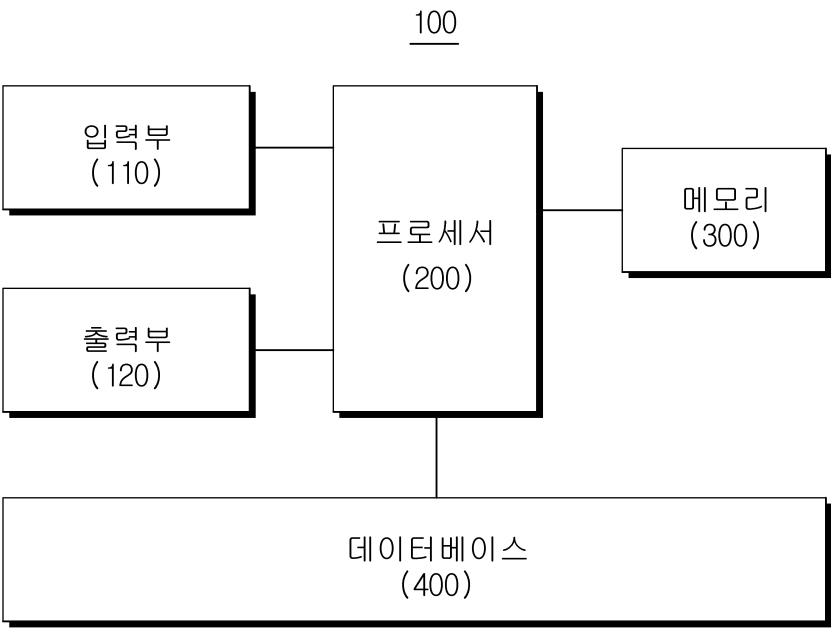
[0093]	100: 질병 진단 장치	110: 입력부
	120: 출력부	200: 프로세서
	300: 메모리	400: 데이터 베이스
	210: 데이터 획득부	220: 데이터 전처리부
	230: 데이터 학습부	240: 질병 진단부

도면

도면1

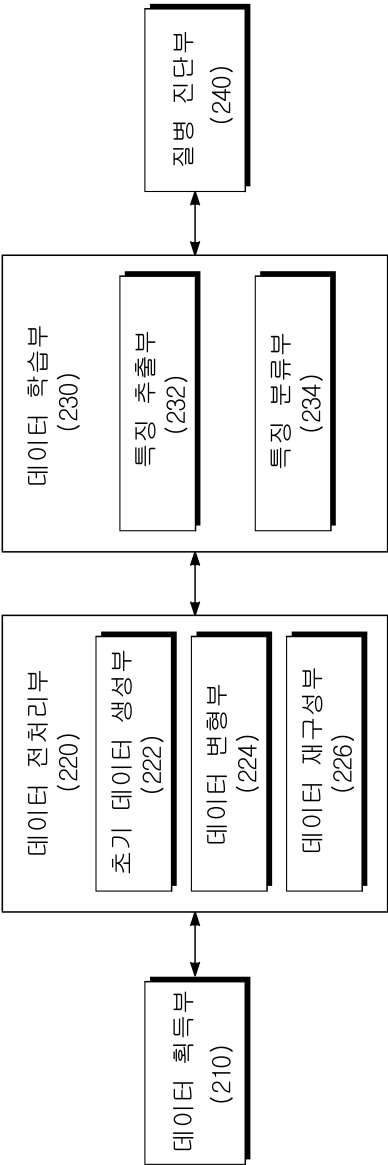


도면2

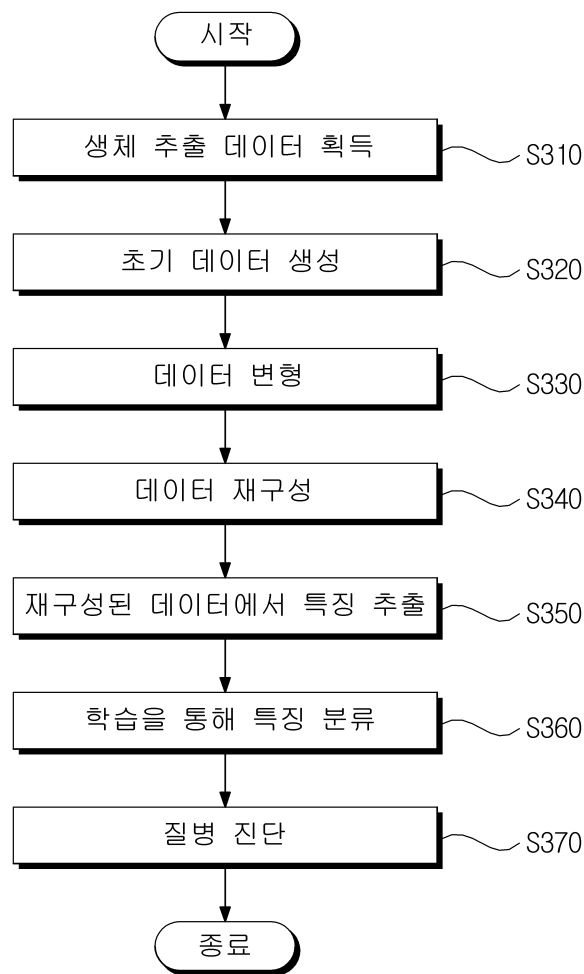


도면3

200

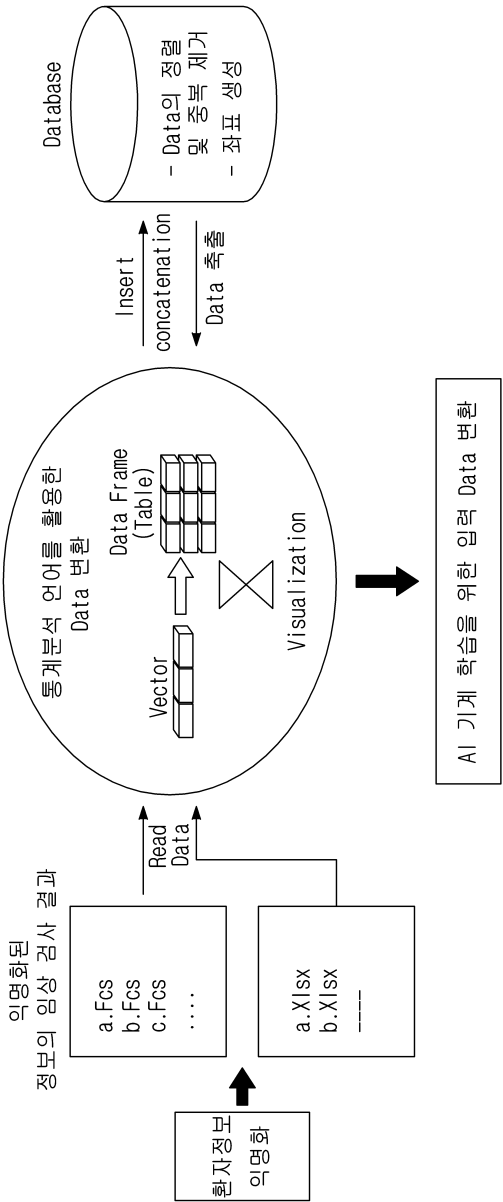


도면4



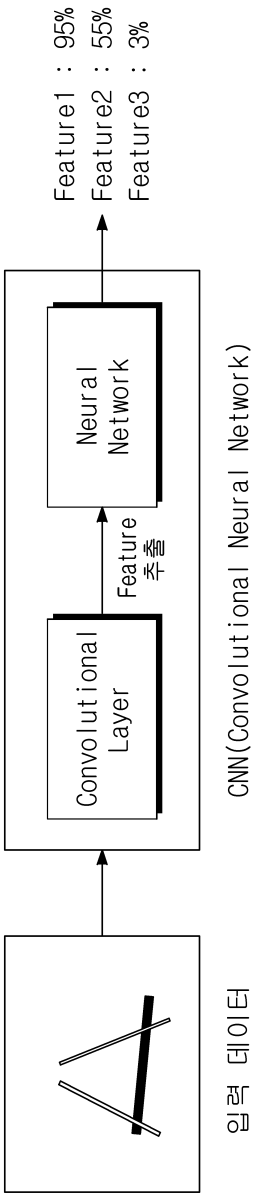
도면5

220



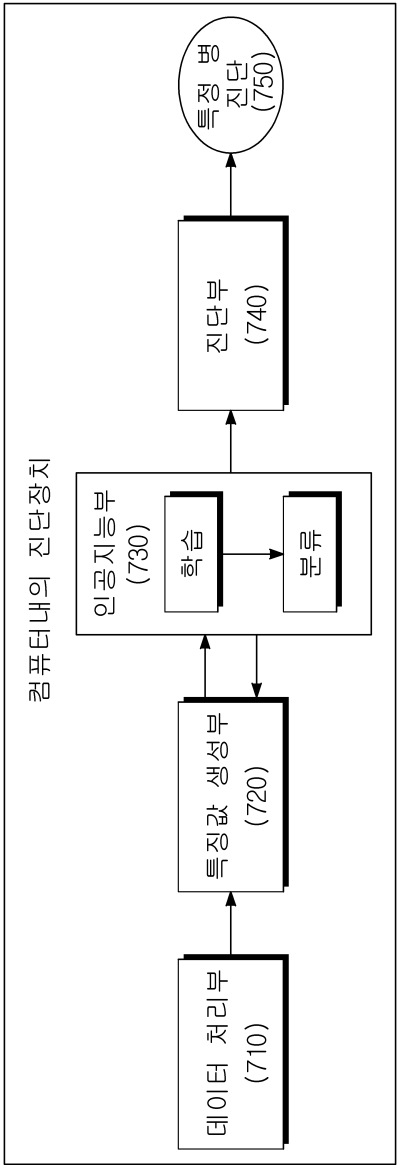
도면6

230



도면7

700

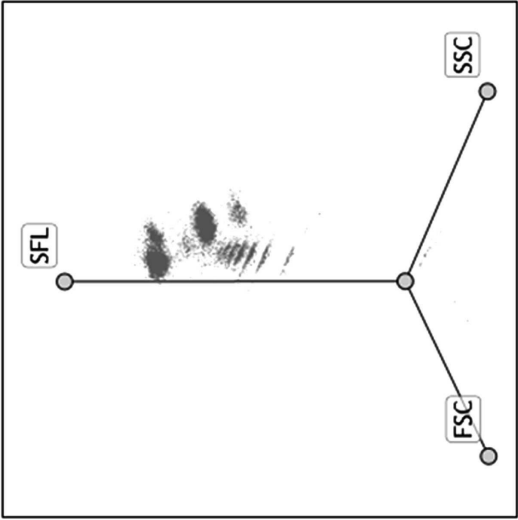


도면8

FILE NAME "*****.fcs"
SERIAL NUMBER "11-222-333"
DATE "yy-mm-dd-hh"
Par1 Par2 Par3 Par4
"FSC INT LIN" "SSC LIN" "FSCW LIN" "SFL LOG"
256 256 256 Resolution

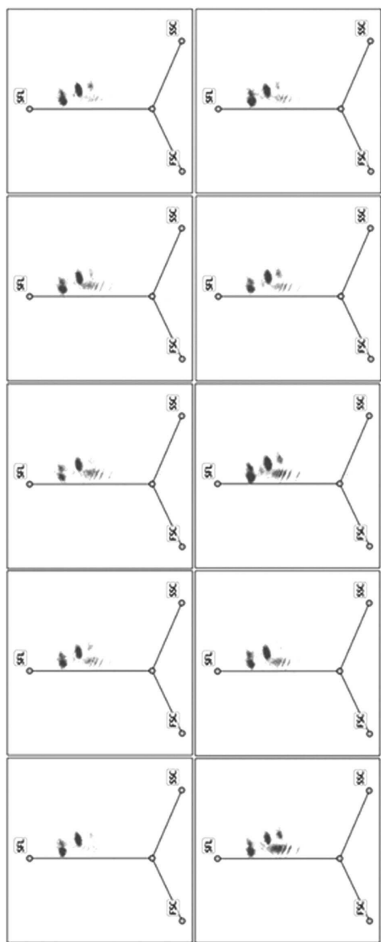
Par1	Par2	Par3	Par4
76	39	32	52
229	93	23	44
182	77	31	36
78	15	39	49
205	76	44	37
47	24	51	36
21	17	26	42
152	56	39	27
45	14	17	28
95	51	39	47
155	85	42	35
149	56	35	28
205	71	29	45
78	37	37	32
22	16	41	43
177	107	41	22
28	16	35	38
125	51	56	49
36	20	50	27
:	:	:	:

(a)

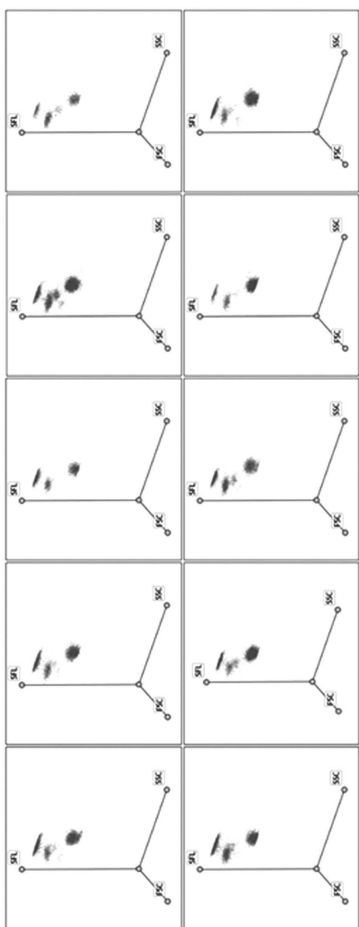


(b)

도면9a

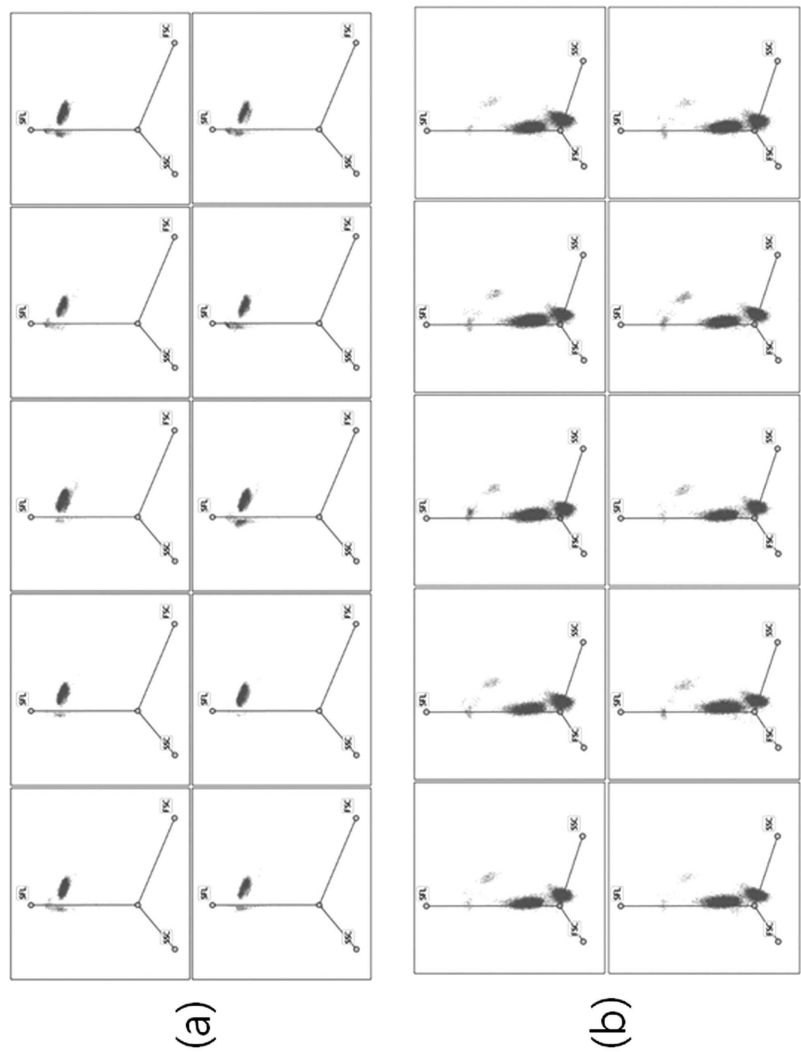


(a)



(b)

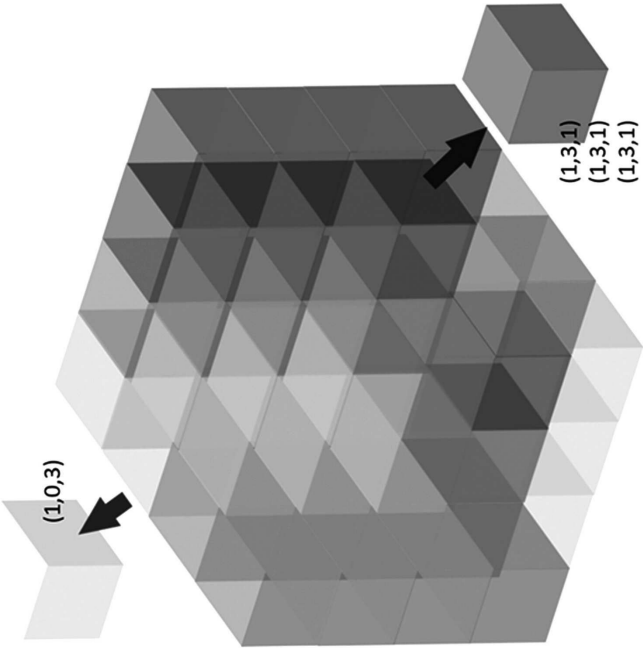
도면9b



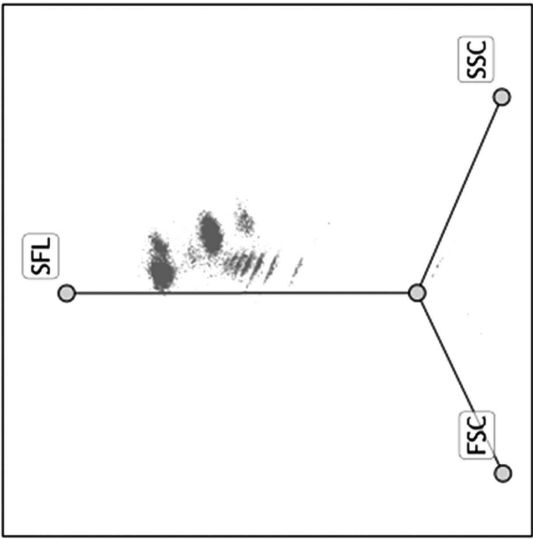
도면10a

Par1Par2Par3	Count
000000013	1
(좌표를 나타냄)	
⋮	⋮
125205115	1
125205116	2
⋮	⋮
034618175	2
⋮	⋮

도면10b



(b)



(a)

도면11

0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	2	0	0	0	0	0	0	0
0	0	0	3	0	0	0	0	0	0	0	0
0	0	0	3	0	0	0	0	0	0	0	0
0	0	0	4	0	0	0	0	0	0	0	0
0	0	0	2	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0