



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0157086
(43) 공개일자 2021년12월28일

(51) 국제특허분류(Int. Cl.)
G06T 7/277 (2017.01) G06N 3/08 (2006.01)
G06T 7/246 (2017.01)
(52) CPC특허분류
G06T 7/277 (2017.01)
G06N 3/08 (2013.01)
(21) 출원번호 10-2020-0074933
(22) 출원일자 2020년06월19일
심사청구일자 2020년06월19일
기술이전 희망 : 기술양도

(71) 출원인
인천대학교 산학협력단
인천광역시 연수구 아카데미로 119 (송도동)
(72) 발명자
최욱
인천광역시 연수구 송도문화로28번길 28, 106동
2802호(송도동, 송도글로벌캠퍼스푸르지오)
이민철
인천광역시 연수구 컨벤시아대로274번길 15,
2304(송도동, 송도 더샵 마스터뷰 23-1BL)
(74) 대리인
특허법인충정

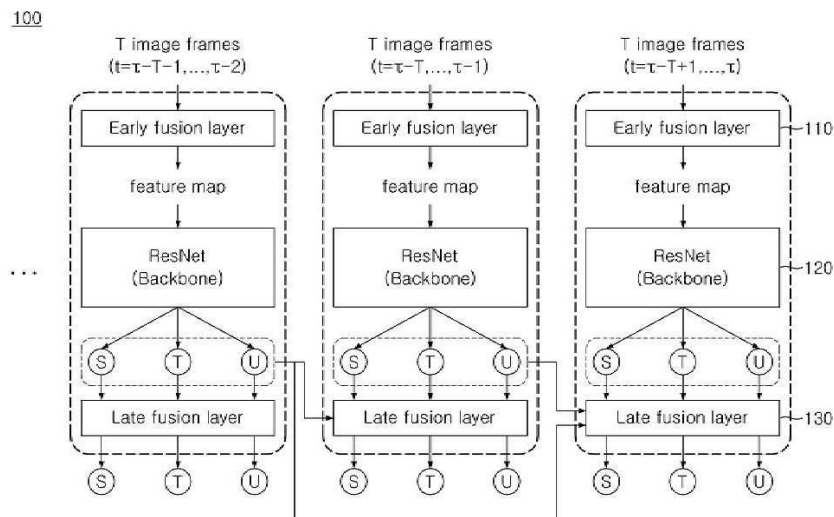
전체 청구항 수 : 총 7 항

(54) 발명의 명칭 인공지능망의 딥러닝을 이용하는 영상 시퀀스 분류 방법

(57) 요약

본 발명은 인공지능망의 딥러닝을 이용하는 영상 시퀀스 분류 방법에 관한 것으로서, 본 발명의 영상 처리 시스템에서 불꽃의 연소 상태를 예측하기 위한 영상 시퀀스 분류 방법은, 전기 융합 계층에서, 불꽃이 포함된 영상 프레임으로서, 복수의 시간 단계에 대하여 각각 T개(T는 1 이상의 자연수)의 순차적인 영상 프레임을 입력받고, 해당 특징 맵 정보를 생성하는 단계, 백본망에서, 각 시간 단계에 대한 상기 특징 맵 정보에 대하여 연소 상태별 확률을 계산하는 단계, 및 후기 융합 계층에서, 상기 각 시간 단계에 대한 연소 상태별 확률로부터 해당 마지막 현재 시간 단계의 연소 상태를 결정하는 단계를 포함한다.

대표도



(52) CPC특허분류

G06T 7/246 (2017.01)

G06T 2207/20084 (2013.01)

G06T 2207/20221 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711116148
과제번호	IITP-2020-2018-0-01431
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	대학ICT연구센터지원사업
연구과제명	MR-IoT융합기반의 재난대응 인공지능 응용기술
기 여 율	1/1
과제수행기관명	아주대학교 산학협력단
연구기간	2018.06.01 ~ 2021.12.31

명세서

청구범위

청구항 1

영상 처리 시스템에서 불꽃의 연소 상태를 예측하기 위한 영상 시퀀스 분류 방법에 있어서,

전기 융합 계층에서, 불꽃이 포함된 영상 프레임으로서, 복수의 시간 단계에 대하여 각각 T개(T는 1 이상의 자연수)의 순차적인 영상 프레임을 입력받고, 해당 특징 맵 정보를 생성하는 단계;

백본망에서, 각 시간 단계에 대한 상기 특징 맵 정보에 대하여 연소 상태별 확률을 계산하는 단계; 및

후기 융합 계층에서, 상기 각 시간 단계에 대한 연소 상태별 확률로부터 해당 마지막 현재 시간 단계의 연소 상태를 결정하는 단계

를 포함하는 영상 시퀀스 분류 방법.

청구항 2

제1항에 있어서,

상기 영상 시퀀스 분류 방법은, 연소 장치에서 캡처된 상기 불꽃이 포함된 영상 프레임에 대하여, 상기 연소 장치의 현재의 불꽃에 대한 상기 연소 상태로서, 안정(S), 과도(T), 또는 불안정(U) 상태 중 어느 상태인지 여부를 예측하기 위한 영상 시퀀스 분류 방법.

청구항 3

제1항에 있어서,

상기 전기 융합 계층은, 상기 순차적인 영상 프레임의 각 픽셀의 휘도값에 대한 주파수 도메인에서의 실수부(Re)와 허수부(Im), 해당 파워 스펙트럼 밀도, 및 상기 순차적인 영상 프레임을 포함한, 상기 특징 맵 정보를 제공하는 영상 시퀀스 분류 방법.

청구항 4

제3항에 있어서,

상기 전기 융합 계층은, 상기 백본망에서 상기 특징 맵 정보에 대한 학습을 통해 결정된 상기 실수부(Re)와 허수부(Im)에 기초한 상기 특징 맵 정보를 제공하는 영상 시퀀스 분류 방법.

청구항 5

제1항에 있어서,

상기 백본망은, 컨볼루션 인공신경망을 이용하여 상기 특징 맵 정보에 대한 학습을 통해 상기 연소 상태별 확률을 계산하는 영상 시퀀스 분류 방법.

청구항 6

제1항에 있어서,

상기 후기 융합 계층은, 상기 연소 상태로서, 안정(S), 과도(T), 불안정(U) 상태 각각에 대한 확률의 3T 차원 벡터 $\hat{y}^{(n)}$ 를 산출하고, 하기의 수학식에 따라 연소 상태별 확률 벡터 $\tilde{y}^{(\tau)}(s)$ 를 산출하는,

$$\tilde{y}^{(\tau)}(s) = \sum_{n=0}^{3T-1} w(s, n) \cdot \hat{y}^{(n)}$$

$$w(s, n) = \begin{cases} 1, & \text{if } n = 3q + s \text{ for } q = 0, \dots, T-1 \\ 0, & \text{otherwise,} \end{cases}$$

영상 시퀀스 분류 방법.

청구항 7

불꽃의 연소 상태를 예측하기 위한 영상 시퀀스 분류를 위한 영상 처리 시스템에 있어서,

불꽃이 포함된 영상 프레임으로서, 복수의 시간 단계에 대하여 각각 T개(T는 1 이상의 자연수)의 순차적인 영상 프레임을 입력받고, 해당 특징 맵 정보를 생성하는 전기 융합 계층;

각 시간 단계에 대한 상기 특징 맵 정보에 대하여 연소 상태별 확률을 계산하는 백본망; 및

상기 각 시간 단계에 대한 연소 상태별 확률로부터 해당 마지막 현재 시간 단계의 연소 상태를 결정하는 후기 융합 계층

을 포함하는 영상 처리 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 영상 시퀀스 분류 방법에 관한 것으로서, 특히, 인공신경망의 딥러닝을 이용하는 영상 시퀀스 분류 방법에 관한 것이다.

배경 기술

[0002] 가스 터빈 등 연소 장치의 연소 불안정성은 동적 압력 및 열 방출에 의한 큰 진동을 일으켜 터빈 자체 및 그에 부착된 터빈 블레이드 및 날개와 같은 고온 가스 경로 상의 구성 요소를 손상시킬 수 있다. 따라서, 이와 같은 연소 불안정성은 가스 터빈이나 내부 엔진을 포함한 많은 연소 장치의 주요 이슈 중 하나였다. 연소 불안정성 메커니즘을 알리고 특성화하며 연소 불안정성을 모니터링하고 제어하기 위해 엄청난 노력을 기울였지만, NOx 배출 규제에 강력한 강화로 인해 가스 터빈이 저배합 사전 혼합 연소 조건에서 작동하게 되었다. 분출 한계 부근의 이러한 사전 혼합 연소는 불꽃을 진동시켜서 연소 불안정성을 증폭시킴으로써 가스 터빈의 안전한 작동을 어렵게 만든다. 그러나 4 세대 혁명의 시대에 고급 머신 러닝 기술의 출현과 함께 연소 불안정성 모니터링 기술의 새로운 패러다임이 적용되고 있다.

[0003] 특히, 광학 감지 및 디지털 이미지 처리 방법이 발전함에 따라 디지털 영상 장치는 온라인 연소 상태 모니터링 시스템에 사용되어왔다. 예를 들어, Lu et al., Sun et al., Bai et al. 등은 불꽃 영상 처리 기술을 제안하였다. 디지털 영상 처리 기술의 발전은 중요한 기능을 자동으로 학습할 수 있는 기계 학습 방법과 결합하여 가속화되고 있다. 따라서, 딥러닝 기반의 머신 러닝 방법은 최근 다양한 컴퓨터 비전 애플리케이션을 위한 새로운 기술로서 제안되어왔다. 딥 러닝은 연소 상태 모니터링 및 연소 불안정성 발생 식별에도 사용되었다. 퍼니스 연소 상태를 모니터링하고 열 방출률을 예측하기 위해 CNN(Convolutional Neural Network)을 제안한 경우도 있다.

[0004] 그러나, 다양한 머신 러닝 방법이 정확도 측면에서 다소 개선할 부분이 많으므로, 연소 불안정성을 모니터링하기 위한 영상 시퀀스 분류 방법의 정확도를 향상시킬 수 있는 개선된 방법이 요구되고 있다.

발명의 내용

해결하려는 과제

[0005] 따라서, 본 발명은 상술한 문제점을 해결하기 위하여 안출된 것으로, 본 발명의 목적은, 가스 터빈과 같은 연소 장치의 연소 불안정성을 모니터링하는 데 있어서 고속 불꽃 영상의 시퀀스에 대한 인공신경망의 딥러닝을 이용해 정확도 성능을 향상시킬 수 있는 영상 시퀀스 분류 방법을 제공하는 데 있다.

과제의 해결 수단

[0006] 먼저, 본 발명의 특징을 요약하면, 상기의 목적을 달성하기 위한 본 발명의 일면에 따른 영상 처리 시스템에서

불꽃의 연소 상태를 예측하기 위한 영상 시퀀스 분류 방법은, 전기 융합 계층에서, 불꽃이 포함된 영상 프레임으로서, 복수의 시간 단계에 대하여 각각 T개(T는 1 이상의 자연수)의 순차적인 영상 프레임을 입력받고, 해당 특징 맵 정보를 생성하는 단계; 백본망에서, 각 시간 단계에 대한 상기 특징 맵 정보에 대하여 연소 상태별 확률을 계산하는 단계; 및 후기 융합 계층에서, 상기 각 시간 단계에 대한 연소 상태별 확률로부터 해당 마지막 현재 시간 단계의 연소 상태를 결정하는 단계를 포함한다.

[0007] 상기 영상 시퀀스 분류 방법은, 연소 장치에서 캡처된 상기 불꽃이 포함된 영상 프레임에 대하여, 상기 연소 장치의 현재의 불꽃에 대한 상기 연소 상태로서, 안정(S), 과도(T), 또는 불안정(U) 상태 중 어느 상태인지 여부를 예측하기 위한 것을 특징으로 한다.

[0008] 상기 전기 융합 계층은, 상기 순차적인 영상 프레임의 각 픽셀의 휘도값에 대한 주파수 도메인에서의 실수부(Re)와 허수부(Im), 해당 파워 스펙트럼 밀도, 및 상기 순차적인 영상 프레임을 포함한, 상기 특징 맵 정보를 제공한다.

[0009] 상기 전기 융합 계층은, 상기 백본망에서 상기 특징 맵 정보에 대한 학습을 통해 결정된 상기 실수부(Re)와 허수부(Im)에 기초한 상기 특징 맵 정보를 제공할 수 있다.

[0010] 상기 백본망은, 컨볼루션 인공신경망을 이용하여 상기 특징 맵 정보에 대한 학습을 통해 상기 연소 상태별 확률을 계산한다.

[0011] 상기 후기 융합 계층은, 상기 연소 상태로서, 안정(S), 과도(T), 불안정(U) 상태 각각에 대한 확률의 3T 차원 벡터 $\hat{y}^{(n)}$ 를 산출하고, 하기의 수학적식에 따라 연소 상태별 확률 벡터 $\hat{y}^{(\tau)}(s)$ 를 산출한다.

$$\hat{y}^{(\tau)}(s) = \sum_{n=0}^{3T-1} w(s, n) \cdot \hat{y}(n)$$

$$w(s, n) = \begin{cases} 1, & \text{if } n = 3q + s \text{ for } q = 0, \dots, T-1 \\ 0, & \text{otherwise,} \end{cases}$$

[0014] 그리고, 본 발명의 다른 일면에 따른 불꽃의 연소 상태를 예측하기 위한 영상 시퀀스 분류를 위한 영상 처리 시스템은, 불꽃이 포함된 영상 프레임으로서, 복수의 시간 단계에 대하여 각각 T개(T는 1 이상의 자연수)의 순차적인 영상 프레임을 입력받고, 해당 특징 맵 정보를 생성하는 전기 융합 계층; 각 시간 단계에 대한 상기 특징 맵 정보에 대하여 연소 상태별 확률을 계산하는 백본망; 및 상기 각 시간 단계에 대한 연소 상태별 확률로부터 해당 마지막 현재 시간 단계의 연소 상태를 결정하는 후기 융합 계층을 포함한다.

발명의 효과

[0015] 본 발명에 따른 영상 시퀀스 분류 방법에 따르면, 기존과 같은 단일 영상의 처리가 아닌 고속 불꽃 영상의 시퀀스에 대한 인공신경망의 딥러닝을 이용한 처리에 따라 가스 터빈과 같은 연소 장치의 연소 불안정성을 모니터링하는 데 있어서 연소 상태 여부를 좀 더 정확하게 파악할 수 있다.

[0016] 본 발명에서는 백본망과 결합하여 성능을 향상시킬 수 있는 두 개의 서로 다른 네트워크 계층으로서, 1) 특정 영상 조건에서 시간에 따라 변하지 않는 순차 영상 프레임들의 파워 스펙트럼 밀도를 추출하는 방법을 적용한 전기 융합 계층, 2) 인접 시간 단계에 걸친 전이 확률을 사용하여 불꽃의 복잡한 시간적 거동을 모델링하는 기존의 방법 대신에, 넓은 범위의 출력을 결합하되 서로 다른 시간에서의 백본망의 출력을 효과적으로 결합하여 현재의 연소 상태를 예측하도록 영상의 분류 정확도를 최대화할 수 있는 후기 융합 계층을 이용해, 계층간 연동 처리로부터 다양한 종류의 연소 불안정성에 대한 예측의 견고성을 보장할 수 있다.

도면의 간단한 설명

[0017] 본 발명에 관한 이해를 돕기 위해 상세한 설명의 일부로 포함되는 첨부도면은, 본 발명에 대한 실시예를 제공하고 상세한 설명과 함께 본 발명의 기술적 사상을 설명한다.

도 1은 본 발명의 영상 처리 시스템에서 시간에 따른 전기 융합 계층, 백본망, 후기 융합 계층 간의 영상 프레임의 처리에 대한 개념도이다.

도 2는 본 발명의 영상 처리 시스템에서의 백본망의 기본 구성의 일례이다.

도 3은 본 발명의 영상 처리 시스템에서 백본망의 구현예이다.

도 4는 본 발명의 영상 처리 시스템에서 전기 융합 계층의 구현예이다.

도 5는 본 발명의 영상 처리 시스템에서 후기 융합 계층의 구현예이다.

도 6은 본 발명의 검증에 사용된 시스템의 가스 터빈 연소 장치 모델과 동적 압력 센서 및 고속 CMOS 디지털 카메라를 보여준다.

도 7은 도 6의 시스템에서 획득된 불꽃에 대한 영상 프레임들의 다양한 경우를 보여준다.

도 8은 본 발명의 일 실시예에 따른 영상 처리 시스템의 구현 방법의 일례를 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0018] 이하에서는 첨부된 도면들을 참조하여 본 발명에 대해서 자세히 설명한다. 이때, 각각의 도면에서 동일한 구성 요소는 가능한 동일한 부호로 나타낸다. 또한, 이미 공지된 기능 및/또는 구성에 대한 상세한 설명은 생략한다. 이하에 개시된 내용은, 다양한 실시 예에 따른 동작을 이해하는데 필요한 부분을 중점적으로 설명하며, 그 설명의 요지를 흐릴 수 있는 요소들에 대한 설명은 생략한다. 또한 도면의 일부 구성요소는 과장되거나 생략되거나 또는 개략적으로 도시될 수 있다. 각 구성요소의 크기는 실제 크기를 전적으로 반영하는 것이 아니며, 따라서 각각의 도면에 그려진 구성요소들의 상대적인 크기나 간격에 의해 여기에 기재되는 내용들이 제한되는 것은 아니다.
- [0019] 본 발명의 실시예들을 설명함에 있어서, 본 발명과 관련된 공지기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하기로 한다. 그리고, 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다. 상세한 설명에서 사용되는 용어는 단지 본 발명의 실시 예들을 기술하기 위한 것이며, 결코 제한적이어서는 안 된다. 명확하게 달리 사용되지 않는 한, 단수 형태의 표현은 복수 형태의 의미를 포함한다. 본 설명에서, "포함" 또는 "구비"와 같은 표현은 어떤 특성들, 숫자들, 단계들, 동작들, 요소들, 이들의 일부 또는 조합을 가리키기 위한 것이며, 기술된 것 이외에 하나 또는 그 이상의 다른 특성, 숫자, 단계, 동작, 요소, 이들의 일부 또는 조합의 존재 또는 가능성을 배제하도록 해석되어서는 안 된다.
- [0020] 또한, 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되는 것은 아니며, 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.
- [0021] 도 1은 본 발명의 영상 처리 시스템(100)에서 시간에 따른 전기 융합 계층(early-fusion layer)(110), 백본망(backbone)(120), 후기 융합 계층(late-fusion layer)(130) 간의 영상 프레임의 처리에 대한 개념도이다. 여기서, "S"(stable, 안정), "T"(transient, 과도), "U"(unstable, 불안정)는 3가지 연소 상태를 나타낸다. 안정(S)은 연소가 안정된 상태, 불안정(U)은 연소가 불안정된 상태, 과도(T)는 안정(S)과 불안정(U) 사이의 상태에서 해당 연소 장치(도 6 참조)의 연소 챔버 내부가 미리 정해진 압력 범위에 있는 경우에 해당한다.
- [0022] 본 발명의 영상 처리 시스템(100)에서는 순차 불꽃 영상 프레임들(연소 장치에서 캡처된 불꽃이 포함된 영상 프레임들)로부터 해당 마지막 프레임에서의 불꽃의 3가지 연소 상태(예, 안정(S), 과도(T), 불안정(U)) 중 어느 상태인지를, 영상 시퀀스에 대해 파악해 나가는, 영상 시퀀스 분류를 수행한다.
- [0023] 본 발명의 영상 처리 시스템(100)은, 백본망(120)(예, 2015년 ImageNet ISLRC(ImageNet Large Scale Visual Recognition Challenge)에서 우승한 아키텍처인 ResNet)과 결합하여 성능을 향상시킬 수 있는 두 개의 서로 다른 네트워크 계층으로서, 1) 특정 영상 조건에서 시간에 따라 변하지 않는 순차 영상 프레임들의 파워 스펙트럼 밀도를 추출하는 방법을 적용한 전기 융합 계층(110), 2) 인접 시간 단계에 걸친 전이 확률을 사용하여 불꽃의 복잡한 시간적 거동을 모델링하는 기존의 방법 대신에, 넓은 범위의 출력을 결합하되 서로 다른 시간에서의 백본망의 출력을 효과적으로 결합하여 현재의 연소 상태를 예측하도록 영상의 분류 정확도를 최대화할 수 있는 후기 융합 계층(130)을 이용한다. 전기 융합 계층(110)은 영상 프레임으로부터 추출된 입력 특징을 백본망(120)에 제공하여 백본망(120)이 더 깊은 특징을 추출하고 연소 상태를 더 잘 예측할 수 있게 한다. 후기 융합 계층(130)은 이전 출력이 현재 출력을 개선하는 데 도움이 될 수 있다는 가정에 따라 백본망(120) 출력들을 서로 다

른 시간 단계에서 결합한다.

[0024] 도 1에서, 전기 융합 계층(early-fusion layer)은, 복수의 시간 단계에서 각각 T개의 순차적인 영상 프레임의 시퀀스를 입력받으며 특징 맵 정보(예, 4T 채널)를 출력한다. T는 1 이상의 자연수이며 바람직하게는 1 내지 16 중 어느 하나일 수 있다. 여기서, 현재 시간 단계 (τ)에서의 영상 프레임들 및 이전 시간 단계($\tau-1$, $\tau-2, \dots$)에서의 각각의 영상 프레임들에 대하여 각각 특징 맵 정보를 생성하여 출력한다.

[0025] 백본망(backbone)은 각 시간 단계에 대한 특징 맵 정보를 기초로 3가지 연소 상태("S" 안정, "T"과도, "U" 불안정 상태)에 대한 확률을 제공한다. 여기서, 현재 시간 단계 (τ)에서의 확률 및 이전 시간 단계($\tau-1$, $\tau-2, \dots$)에서의 각각의 확률이 제공된다.

[0026] 후기 융합 계층(late-fusion layer)은 백본망(backbone)으로부터의 현재 시간 단계 (τ) 및 이전 시간 단계($\tau-1$, $\tau-2, \dots$)에서의 3가지 연소 상태에 대한 확률들로부터 해당 마지막 현재 시간 단계 (τ)에서의 연소 상태를 결정한다.

[0028] 한편, 본 발명의 백본망(120)에서의 신경망 학습을 위한 컨볼루션 연산을 먼저, 간략히 설명한다.

[0030] < 백본망(120)에서의 신경망 학습을 위한 컨볼루션 연산>

[0031] 영상은 2D 픽셀 격자로 표현되므로, 2D 컨볼루션은 모든 픽셀에 대한 선형 필터(예, $1 \times 1 \sim 7 \times 7$)를 적용한 영상으로 표현될 수 있으며, 예지, 코너, 아핀 공변량 영역(affine covariant regions)과 같은 영상 특징을 추출하는데 이용된다. [수학식1]과 같이 2D 컨볼루션이 C 채널을 갖는 영상 I에 대하여 연산되어 픽셀 좌표(i, j)에 대한 필터 응답 H_F 이 특징 맵 정보로서 출력될 수 있다. 여기서, c는 채널 인덱스, $I(c, i+u, j+v)$ 는 채널 c에서 ($i+u, j+v$)의 픽셀값(휘도값), $(2U+1) \times (2V+1)$ 은 필터 사이즈, $W(c, u, v)$ 는 필터 가중치, b는 바이어스값으로서 0이상의 실수이다. $W(c, u, v)$ 와 b는 딥러닝에 의해 학습된다.

[0032] [수학식1]

$$H_F(i, j) = \sum_{c=0}^{C-1} \sum_{u=-U}^U \sum_{v=-V}^V W(c, u, v) \cdot I(c, i+u, j+v) + b$$

[0033]

[0034] 특징 맵 정보 H_F 를 획득하기 위하여, [수학식2]와 같이 K(자연수)개의 필터가 사용될 수도 있다. 여기서, k는 0에서 K-1까지의 필터 인덱스.

[0035] [수학식2]

$$H_F(k, i, j) = \sum_{c=0}^{C-1} \sum_{u=-U}^U \sum_{v=-V}^V W(k, c, u, v) \cdot I(c, i+u, j+v) + b(k)$$

[0036]

[0037] 불필요한 특징이 상기 필터에 의해 제거되도록 최적의 필터가 사용되는 맥스-풀링(max-pooling) 방식이 적용되며, 이는 메모리 사이즈를 줄이고 딥러닝의 정확도를 높여준다.

[0038] 인공신경망의 딥러닝은 적절한 학습율(learning rate)에 의해 적절한 훈련 시간 동안 이루어지도록 하며, 특징 맵 정보 H_F 는 [수학식3]과 같이 평균 $E[\]$ 와 분산 $\text{Var}[\]$ 을 사용해 정규화(normalizing)/백색화(whitening)된 값 H_N 으로 표현될 수 있다. 여기서, ϵ 은 분모가 0되지 않도록 하기 위한 소장의 실수이고, H_N 의 평균과 분산은 각각 0과 1이 된다.

[0039] [수학식3]

$$H_N(k, i, j) = \frac{H_F(k, i, j) - E[H_F(k)]}{\sqrt{\epsilon + \text{Var}[H_F(k)]}}$$

[0040]

[0041] 이와 같은 정규화에 의한 representation power의 손실을 복구하기 위하여, [수학식4]와 같이 학습 가능한 계수 γ , β 를 이용한 특징 맵 정보의 선형 변환값 H_{BN} 이 이용될 수 있다. 여기서, β 는 바이어스값 b 또는 $b(k)$ 의 역할을 할 수 있다.

- [0042] [수학식4]
- [0043]
$$H_{BN}(k,i,j) = \gamma \cdot H_N(k,i,j) + \beta$$
- [0044] 영상에 상기 선형 필터들을 반복 적용하는 것은 복합 선형 필터를 적용하는 것과 같은 효과를 가지며, 비선형성을 갖도록 하기 위하여 [수학식5]와 같은 정류 선형 유닛(Rectified linear unit) $H(k,i,j)$ 을 정의해 사용함으로써 인공신경망의 텡쓰(depth)를 증가시킬 수 있다.
- [0045] [수학식5]
- [0046]
$$H(k,i,j) = \max(0, H_{BN}(k,i,j))$$
- [0048] <백본망(120) (예, ResNet)>
- [0049] 컨볼루션 인공신경망의 순차 계층들의 수(또는 텡쓰)는, 성능 향상에 중요하지만, 훈련 에러를 증가시킬 수는 있다. 따라서, 왜곡을 줄이기 위하여 적절한 수의 인공신경망의 계층들이 빠른 연산이 이루어질 수 있는 연결 형태로 이루어지는 것이 바람직하다.
- [0050] 도 2는 본 발명의 영상 처리 시스템(100)에서의 백본망(120)의 기본 구성의 일례이다. 도 2의 (a)는 기본 백본망(예, ResNet-18)의 컨볼루션 블록(basic block)을 나타내고, 도 2의 (b)는 텡쓰를 증가시킨 백본망(예, ResNet-50)의 컨볼루션 블록(bottleneck block)을 나타낸다. 여기서, m은 필터의 개수, 4m은 m의 4배의 필터 개수, d는 소정의 양의 정수이다. 도 2의 (b)의 컨볼루션 블록(bottleneck block) 구조가 도 2의 (a)의 컨볼루션 블록(basic block) 구조 보다 훨씬 정확도를 높여준다.
- [0051] 백본망(120)의 계층들에서 입력 특징 맵 정보 X를 처리하는 H 함수(정류 선형 유닛(ReLU))에 의해, [수학식6]과 같이 출력 특징 맵 정보 Y가 생성되며, 레지듀얼(residual) 매핑값 $F(X) = H(X) - X$ 가 되도록 학습되어, [수학식7]과 같이 Y는 거의 X와 같게 된다.
- [0052] [수학식6]
- [0053]
$$Y = \mathcal{H}(X) = \mathcal{F}(X) + X$$
- [0054] [수학식7]
- [0055]
$$Y = X$$
- [0056] 도 3은 본 발명의 영상 처리 시스템(100)에서 백본망(120)의 구현예이다. 도 3의 (a)는 도 2의 (a)의 컨볼루션 블록(basic block)을 이용한 구현예이고, 도 3의 (b)는 도 2의 (b)의 컨볼루션 블록(bottleneck block)을 이용한 구현예이다. 여기서, 각 컨볼루션 블록의 우측 숫자는 필터 개수, "/2"는 특징 맵 정보(벡터)가 반수가 되도록 띄엄띄엄 원소를 취하거나(stride=2) 맥스-풀링한 것을 나타낸다.
- [0057] 도 3의 (a), (b)에서, 각 입력 영상 프레임의 픽셀 사이즈는 224X224이며, 먼저, 첫번째 블록에서 입력 영상 프레임에 7X7 컨볼루션 연산을 수행하되 (stride=2)를 취하여 수행해 픽셀 사이즈 112X112의 특징 맵 정보를 생성한다. 다음에 이와 같은 특징 맵 정보는 56X56 픽셀 사이즈로 맥스-풀링될 수 있다. 다음에 도면과 같은 일련의 컨볼루션 블록들에서 처리되어, 최종 특징 맵 정보는 7X7 픽셀 사이즈로 출력될 수 있다. 다음에 최종 특징 맵 정보의 원소들을 평균하여 각 채널에 대한 평균값(avg pool)의 1X1 픽셀 사이즈로 출력될 수 있다. 다음에 다차원(예, 3차원)으로 통합 연결된(fully connected) 계층(fc 3)에서의 연산을 통해 예측된 3가지 연소 상태 $S = \{ "S", "T", "U" \}$ 의 각 상태 s에 대한 확률 $P(s|I)$ 을 출력한다.
- [0058] 이때, 백본망(120)의 출력 벡터 $y(s)$ 각각은, [수학식8]과 같이, 백본망 전체 계층에서의 상태 s에 대한 출력 $o(s)$ 의 softmax 함수로서, $y(s)$ 를 최대가 되게하는 상태 s에 의해 결정되고, 이때의 상태 s에 대한 각각의 확률 $P(s|I)$ 을 나타낸다. 여기서, s는 가지 연소 상태의 집합 $S = \{ "S", "T", "U" \}$ 의 원소이다.

[0059] [수학식8]

$$y(s) = P(s|I) = \text{softmax}(o(s)) = \frac{\exp(o(s))}{\sum_{r \in S} \exp(o(r))}$$

[0062] <전기 융합 계층(110) >

[0063] 각 시간 단계 중 시간 $t=\tau$ 에서, 예를 들어, T개의 채널($c=0,1,\dots,T-1$)(시간 $t=\tau-T+1,\dots,\tau$ 에 대한 채널)을 통해 백본망으로 입력되는 순차 영상 프레임들(픽셀값) $I^{(c+\tau-T+1)}$ 는, [수학식9]와 같이, $J^{(\tau)}(c)$ 로 나타내기로 한다(도 5 참조). 해당 채널에서 입력되는 순차 영상 프레임들 $J^{(\tau)}(c)$ 은 시간 $t=\tau-T+1,\dots,\tau$ 와 같이 T개의 프레임들일 수 있다.

[0064] [수학식9]

$$J^{(\tau)}(c) = I^{(c+\tau-T+1)}$$

[0066] 연소 불안정성이 발생하면, 불꽃 영상들은 시간에 따라 빠르게 변한다. 그 변화는 주기적이며, 이는 가스 터빈과 같은 연소 장치 내의 혼합물의 점성이 주기적으로 진동하거나 연소 챔버 내의 동적 압력의 주기적 변화로부터 기인할 수 있다. 이러한 주기성은 각 픽셀의 휘도값에 대한 이산 푸리에 변환을 이용해 [수학식10]과 같이 주파수 도메인에서 실수부(Re)와 허수부(Im)로 표현될 수 있다. 여기서 k는 주파수 $2\pi kcf/T$ 에 대응되고 주파수이고, f는 영상 샘플링 주파수(예, 8kHz)이다.

[0067] [수학식10]

$$Re(k,i,j) = \sum_{c=0}^{T-1} \cos(2\pi kc/T) \cdot J^{(\tau)}(c,i,j)$$

$$Im(k,i,j) = - \sum_{c=0}^{T-1} \sin(2\pi kc/T) \cdot J^{(\tau)}(c,i,j)$$

[0070] [수학식11]과 같이 k번째 주파수에 대하여 파워 스펙트럼 밀도(PS)를 산출할 수 있으며, 입력 영상 프레임들이 T개씩 입력되는 주기를 가질 때 상기 파워 스펙트럼 밀도는 시간에 따라 변하지 않는 값을 가지며, 이에 따라 백본망에서 좀더 안정되게 학습이 수행되도록 한다.

[0071] [수학식11]

$$PS(k,i,j) = Re^2(k,i,j) + Im^2(k,i,j)$$

[0073] 주기 연소 장치의 연료 성분이나 열 출입과 같은 다양한 요소들에 의해 상기 주기 T가 달라질 수 있으며, 입력 영상 프레임들의 입력때마다 또는 연소 상태에 따라 다를 수 있다. 다만, 본 발명의 하기하는 바와 같은 실험 결과들은 입력 영상 프레임들의 모든 시퀀스에 대하여 T를 고정한 경우이다.

[0074] [수학식10]의 실수부(Re)와 허수부(Im) 각각(F)은 [수학식12]와 같이 다시 표현될 수 있다. 학습 가중치 $W(k,c)$ 는 $c=0,1,\dots,T_{\text{exact}}-1$ 에 대하여 실수부(Re)에서 $\cos 2\pi kc/T_{\text{exact}}$ 이고(허수부(Im)에서 $-\sin 2\pi kc/T_{\text{exact}}$) 그렇지 않으면 0의 값을 가지며, T_{exact} 는 T 보다 작은 미리 결정된 주기이다. T_{exact} 가 입력 영상 프레임들의 입력때마다 달라질 수 있으며, 백본망(120)에서 특징 맵 정보의 학습을 통해 T_{exact} 가 학습 가중치 $W(k,c)$ 와 함께 가장 좋은 값으로 결정될 수 있다. 이때, 전기 융합 계층(110)은 백본망(120)에서 특징 맵 정보에 대한 학습을 통해 결정된 실수부(Re)와 허수부(Im)에 기초한 상기 파워 스펙트럼 밀도 등 특징 맵 정보를 제공한다.

[0075] [수학식12]

$$F(k,i,j) = \sum_{c=0}^{T-1} W(k,c) \cdot J^{(\tau)}(c,i,j)$$

[0077] 도 4는 본 발명의 영상 처리 시스템(100)에서 전기 융합 계층(110)의 구현예이다.

[0078] 도 4와 같이, 입력되는 순차 영상 프레임들 $J^{(\tau)}(c)$ 은, [수학식12]를 연산하기 위한 T개의 필터를 갖는 1X1 콘볼루션 연산을 통해, [수학식10]과 같이 각 픽셀에 대한 이산 푸리에 변환된 실수부(Re)와 허수부(Im) 값으로 변환되며, 각각에 대한 계곱 연산과 더하기 연산을 통해 [수학식11]과 같이 파워 스펙트럼 밀도(PS)를 산출할 수 있다.

[0079] 이와 같은 과정을 통해, 소정의 픽셀 사이즈에 대하여 입력되는 순차 영상 프레임들 $J^{(\tau)}(c)$ 및 해당 실수부(Re), 허수부(Im), 파워 스펙트럼 밀도(PS)를 포함하는 특징 맵 정보가, 상기한 바와 같이 백본망으로 출력되어 학습된다.

[0081] <후기 융합 계층(130) >

[0082] 도 5는 본 발명의 영상 처리 시스템(100)에서 후기 융합 계층(130)의 구현예이다.

[0083] 도 5를 참조하면, 백본망에서 생성되어 출력되는 시간 $t = \tau - T + 1, \dots, \tau$ 에 대한 3가지 연소 상태 $S = \{"S", "T", "U"\}$ 각 상태 s 에 대한 확률 $P(s|I)$ 의 3T 차원 벡터 $\hat{y}^{(n)}$ 를 계산하고, 이는, [수학식13]과 같이 백본망 계층들의 전체 연결 계층에 대응된 가중치 벡터 $w(s, n)$ 에 의해 처리되어, 현재 시간 τ 에서의 3가지 연소 상태에 대한 확률 벡터 $\hat{y}^{(\tau)}(s)$ 을 최종 예측값으로 생성한다. 도면에서 3가지 연소 상태 "S", "T", "U"에 대하여 각각 0, 1, 2로 표시하였다.

[0084] [수학식13]

$$\hat{y}^{(\tau)}(s) = \sum_{n=0}^{3T-1} w(s, n) \cdot \hat{y}^{(n)}$$

[0085] 여기서, $w(s, n)$ 은 미리 결정된 값이 될 수도 있고, 학습에 의해 결정될 수도 있으며, 예를 들어, [수학식14]와 같이 결정될 수도 있다(q 는 순차 영상 프레임의 구분자). [수학식15]는 $T=2$ 일때의 $w(s, n)$ 에 대한 예시이다. [수학식15]에서 행은 s , 열은 n 을 나타낸다.

[0087] [수학식14]

$$w(s, n) = \begin{cases} 1, & \text{if } n = 3q + s \text{ for } q = 0, \dots, T - 1 \\ 0, & \text{otherwise,} \end{cases}$$

[0089] [수학식15]

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

[0091] 이와 같이, 회귀 신경망(RNN, recurrent neural networks)는 T가 커짐에 따라 연산의 복잡성과 학습 시간의 비용이 큰 반면, 본 발명은 이러한 문제점을 해결하여 연산의 복잡도를 줄이고 학습에 걸리는 시간을 훨씬 줄일 수 있다.

[0093] < 모델 검증 테스트 >

[0094] 도 6은 본 발명의 검증에 사용된 시스템의 가스 터빈 연소 장치 모델과 동적 압력 센서 및 고속 CMOS 디지털 카메라를 보여준다.

[0095] 도 6에서, 연료가 가스 터빈 연소 장치의 쿼츠(quartz) 부분에서 연소되어 불꽃을 일으키며 쿼츠 부분(또는 연소 챔버)의 과열을 방지하기 위한 냉각 시스템이 구비될 수 있고, 연소 불안정성은 압력 센서(예, 피에조 센

서)의 센싱값으로 모니터링될 수 있다. 연소 챔버 내에서 연료의 연소 동안 고속 CMOS 디지털 카메라를 통해서 불꽃에 대한 영상 프레임들이 8kHz 단위로 획득되었다.

[0096] 도 7은 도 6의 시스템에서 획득된 불꽃에 대한 영상 프레임들의 다양한 경우를 보여준다. 도면에서 영상 프레임들과 대응되는 압력 센서의 센싱 신호(DP signal)를 대응시켜 나타내었으며, 소정의 임계치를 정하여 3가지 연소 상태 "S"(stable, 안정), "T"(transient, 과도), "U"(unstable, 불안정)에 대하여 서로 다른 색상으로 표시하였고, 각각의 연소 상태에서 하부의 블랙 박스와 화이트 박스 부분이 각각 검증용 프레임과 테스트용 프레임을 나타낸다. 나머지 프레임들은 백본망의 학습에 이용되었다.

[0097] 아래 [표1]은 실험결과 도 2의 (a)의 백본망(예, ResNet-18)을 적용한 경우, $T_e(T_{exact})$ 의 1,2,4,8,16에 대해 종래의 방법(ResNet)과 본 발명의 전기 융합 계층(early-fusion layer)의 파워 스펙트럼 밀도(PS)을 이용하는 방법(PS-ResNet)의 3가지 연소 상태 "S"(stable, 안정), "T"(transient, 과도), "U"(unstable, 불안정)에 대한 분류를 예측하는 정확도(%)의 비교표이다.

[0098] [표1]

Case No.	T_e	1	2	4	8	16
1	ResNet	98.0	99.0	99.0	99.5	98.9
	PS-ResNet	98.7	98.0	98.9	97.4	99.0
2	ResNet	100	100	100	100	100
	PS-ResNet	100	100	100	100	100
3	ResNet	80.7	85.5	90.2	89.6	92.0
	PS-ResNet	86.9	87.7	87.5	84.5	93.0
4	ResNet	100	100	100	100	100
	PS-ResNet	100	100	100	100	100
5	ResNet	96.8	98.6	99.8	100	100
	PS-ResNet	97.4	98.0	99.8	100	100
6	ResNet	99.8	99.8	100	100	100
	PS-ResNet	99.7	100	100	100	100
All	ResNet	95.1	96.5	97.7	97.6	98.1
	PS-ResNet	96.6	96.8	97.1	96.2	98.3

[0099] 아래 [표2]는 실험결과 도 2의 (b)의 백본망(예, ResNet-50)을 적용한 경우, $T_e(T_{exact})$ 의 1,2,4,8,16에 대해 종래의 방법(ResNet)과 본 발명의 전기 융합 계층(early-fusion layer)의 파워 스펙트럼 밀도(PS)을 이용하는 방법(PS-ResNet)의 3가지 연소 상태 "S"(stable, 안정), "T"(transient, 과도), "U"(unstable, 불안정)에 대한 분류를 예측하는 정확도(%)의 비교표이다.

[0101] [표2]

Case No.	T_e	1	2	4	8	16
1	ResNet	96.1	97.5	97.6	99.2	98.8
	PS-ResNet	97.0	96.6	98.2	99.2	99.6
2	ResNet	100	100	100	100	100
	PS-ResNet	100	100	100	100	100
3	ResNet	86.8	79.5	85.9	85.6	89.5
	PS-ResNet	88.1	83.2	86.4	87.3	88.9
4	ResNet	100	100	100	100	100
	PS-ResNet	100	100	100	100	100
5	ResNet	97.0	98.9	99.5	99.8	100
	PS-ResNet	97.7	98.2	99.2	100	100
6	ResNet	99.8	100	100	100	100
	PS-ResNet	100	100	100	100	100
All	ResNet	96.3	95.0	96.5	96.7	97.5
	PS-ResNet	96.8	95.7	96.7	97.1	97.5

[0102] [표1], [표2]와 같이, 종래의 방법(ResNet) 보다 본 발명의 방법(PS-ResNet)은 전체적으로 95.1~98.6%의 높은 정확도(%)를 나타내었으며, 3.5%의 성능 향상이 있음을 확인하였다.

[0103] [표1], [표2]와 같이, 종래의 방법(ResNet) 보다 본 발명의 방법(PS-ResNet)은 전체적으로 95.1~98.6%의 높은 정확도(%)를 나타내었으며, 3.5%의 성능 향상이 있음을 확인하였다.

[0104] 아래 [표3]은 본 발명의 전기 융합 계층(early-fusion layer)이나 후기 융합 계층(late-fusion layer)이 적용되거나, 또는 모두 적용되지 않은 경우, $T_e(T_{\text{exact}})$ 의 1, 16에 대해 전기 융합 계층, 백본망 및 후기 융합 계층에서 처리된 연소 상태 분류를 예측하는 각각의 평균 처리 시간의 비교표이다. 여기서는 소정의 컴퓨터 프로세서(Intel Xeon 4114 프로세서를 탑재한 Ubuntu)와 RAM(예, 256GB)을 이용하여, 연산의 처리 시간이 5.2~12.2ms와 같이 매우 짧은 시간에 처리될 수 있음을 확인하였다.

[0105] [표3]

Backbone	T_e	Early fusion	Late fusion	Time (ms)
ResNet-18	1	no	no	5.21
		yes	no	5.50
		yes	yes	7.88
	16	no	no	7.18
		yes	no	8.33
		yes	yes	8.44
ResNet-50	1	no	no	10.06
		yes	no	10.09
		yes	yes	11.88
	16	no	no	11.49
		yes	no	11.80
		yes	yes	12.17

[0106]

[0107] 도 8은 본 발명의 일 실시예에 따른 영상 처리 시스템(100)의 구현 방법의 일례를 설명하기 위한 도면이다.

[0108] 본 발명의 일 실시예에 따른 영상 처리 시스템(100)(또는 전기 융합 계층(110), 백본망(120), 후기 융합 계층(130))은, 하드웨어, 소프트웨어, 또는 이들의 결합으로 이루어질 수 있다. 예를 들어, 본 발명의 영상 처리 시스템(100)은 위와 같은 기능/단계/과정들을 수행하기 위한 적어도 하나의 프로세서를 갖는 도 8과 같은 컴퓨팅 시스템(1000) 또는 인터넷 상의 서버 형태로 구현될 수 있다.

[0109] 컴퓨팅 시스템(1000)은 버스(1200)를 통해 연결되는 적어도 하나의 프로세서(1100), 메모리(1300), 사용자 인터페이스 입력 장치(1400), 사용자 인터페이스 출력 장치(1500), 스토리지(1600), 및 네트워크 인터페이스(1700)를 포함할 수 있다. 프로세서(1100)는 중앙 처리 장치(CPU) 또는 메모리(1300) 및/또는 스토리지(1600)에 저장된 명령어들에 대한 처리를 실행하는 반도체 장치일 수 있다. 메모리(1300) 및 스토리지(1600)는 다양한 종류의 휘발성 또는 불휘발성 저장 매체를 포함할 수 있다. 예를 들어, 메모리(1300)는 ROM(Read Only Memory)(1310) 및 RAM(Random Access Memory)(1320)을 포함할 수 있다.

[0110] 따라서, 본 명세서에 개시된 실시예들과 관련하여 설명된 방법 또는 알고리즘의 단계는 프로세서(1100)에 의해 실행되는 하드웨어, 소프트웨어 모듈, 또는 그 2 개의 결합으로 직접 구현될 수 있다. 소프트웨어 모듈은 RAM 메모리, 플래시 메모리, ROM 메모리, EPROM 메모리, EEPROM 메모리, 레지스터, 하드 디스크, 착탈형 디스크, CD-ROM과 같이 컴퓨터 등 장치로 판독 가능한 저장/기록 매체(즉, 메모리(1300) 및/또는 스토리지(1600))에 상주할 수도 있다. 예시적인 저장 매체는 프로세서(1100)에 커널링되며, 그 프로세서(1100)는 저장 매체로부터 정보를 판독할 수 있고 저장 매체에 정보를 기입할 수 있다. 다른 방법으로, 저장 매체는 프로세서(1100)와 일체형일 수도 있다. 프로세서 및 저장 매체는 주문형 집적회로(ASIC) 내에 상주할 수도 있다. ASIC는 사용자 단말기 내에 상주할 수도 있다. 다른 방법으로, 프로세서 및 저장 매체는 사용자 단말기 내에 개별 컴포넌트로서 상주할 수도 있다.

[0111] 상술한 바와 같이, 본 발명에 따른 영상 처리 시스템(100)의 영상 시퀀스 분류 방법에 따르면, 기존과 같은 단일 영상의 처리가 아닌 고속 불꽃 영상의 시퀀스에 대한 인공신경망의 딥러닝을 이용한 처리에 따라 가스 터빈과 같은 연소 장치의 연소 불안정성을 모니터링하는 데 있어서 연소 상태 여부를 좀 더 정확하게 파악할 수 있다. 본 발명에서는 백본망과 결합하여 성능을 향상시킬 수 있는 두 개의 서로 다른 네트워크 계층으로서, 1) 특정 영상 조건에서 시간에 따라 변하지 않는 순차 영상 프레임들의 파워 스펙트럼 밀도를 추출하는 방법을 적용한 전기 융합 계층, 2) 인접 시간 단계에 걸친 전이 확률을 사용하여 불꽃의 복잡한 시간적 거동을 모델링하는 기존의 방법 대신에, 넓은 범위의 출력을 결합하되 서로 다른 시간에서의 백본망의 출력을 효과적으로 결합하여 현재의 연소 상태를 예측하도록 영상의 분류 정확도를 최대화할 수 있는 후기 융합 계층을 이용해, 계층간 연동 처리로부터 다양한 종류의 연소 불안정성에 대한 예측의 견고성을 보장할 수 있다.

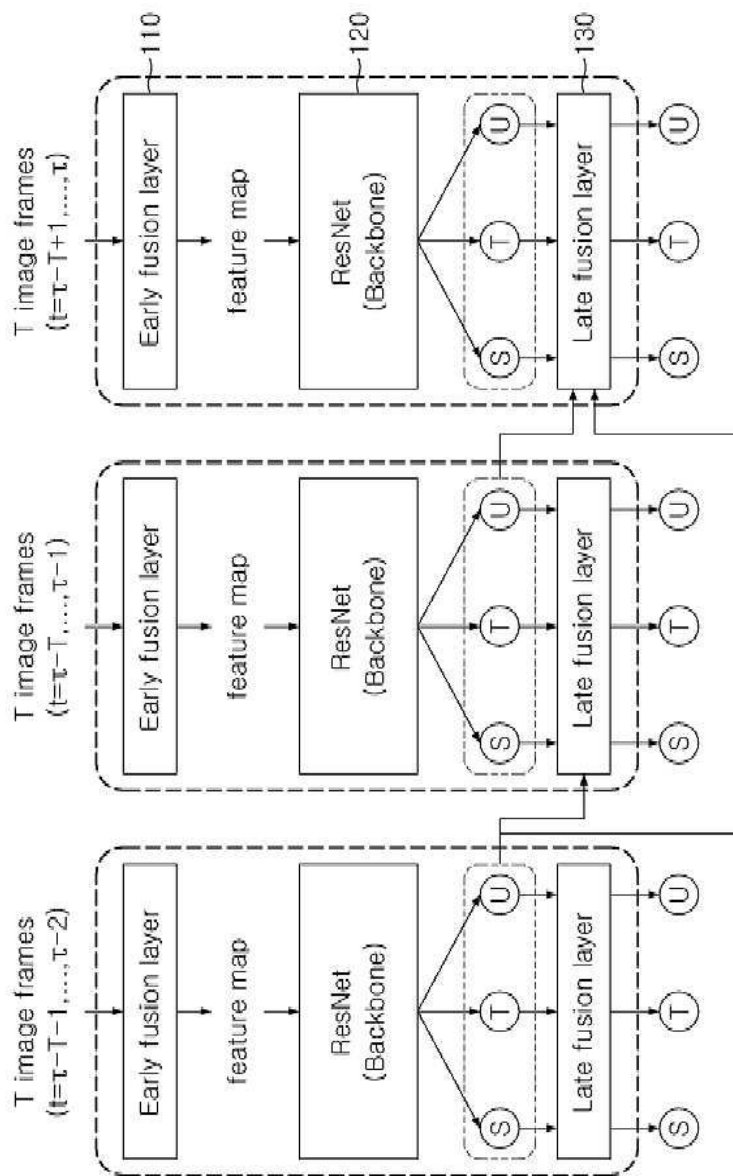
[0112] 이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니 되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

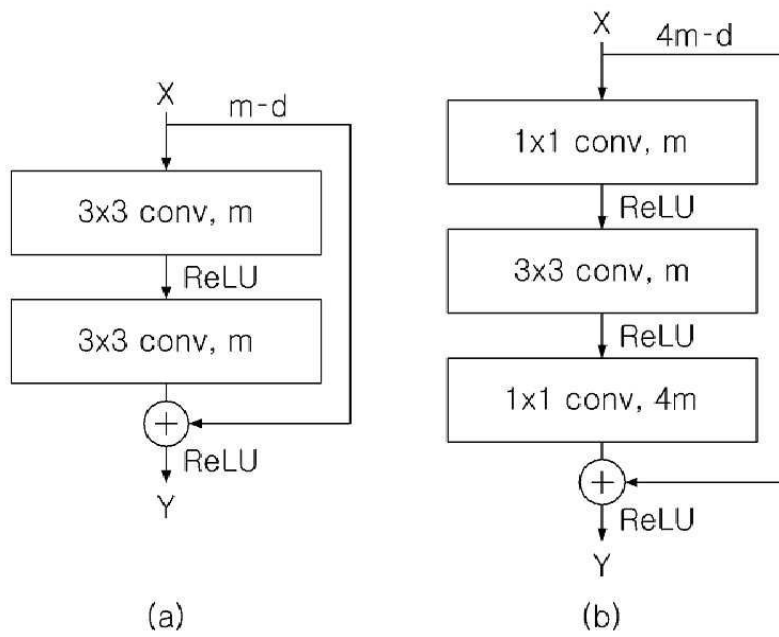
[0113] 전기 융합 계층(110)
 백본망(120)
 후기 융합 계층(130)

도면

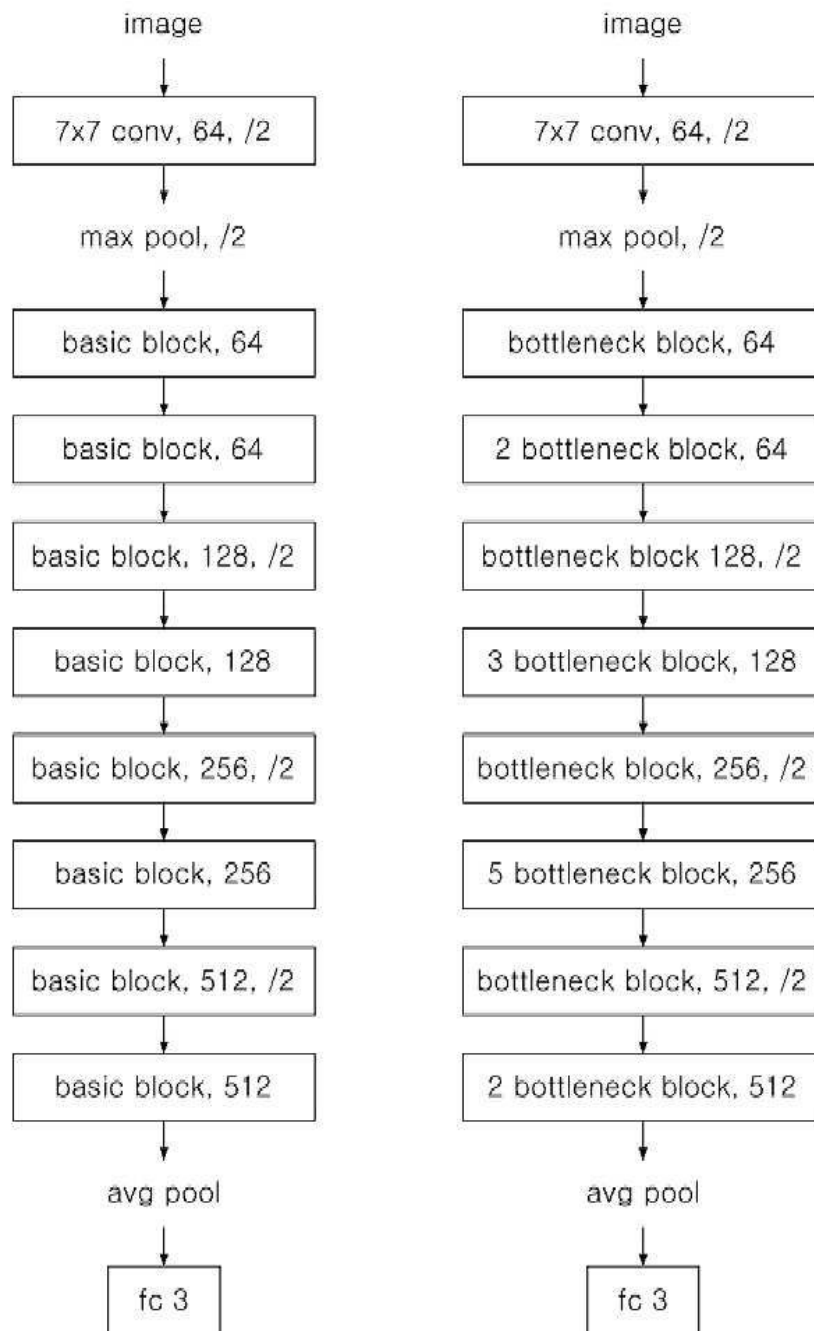
도면1



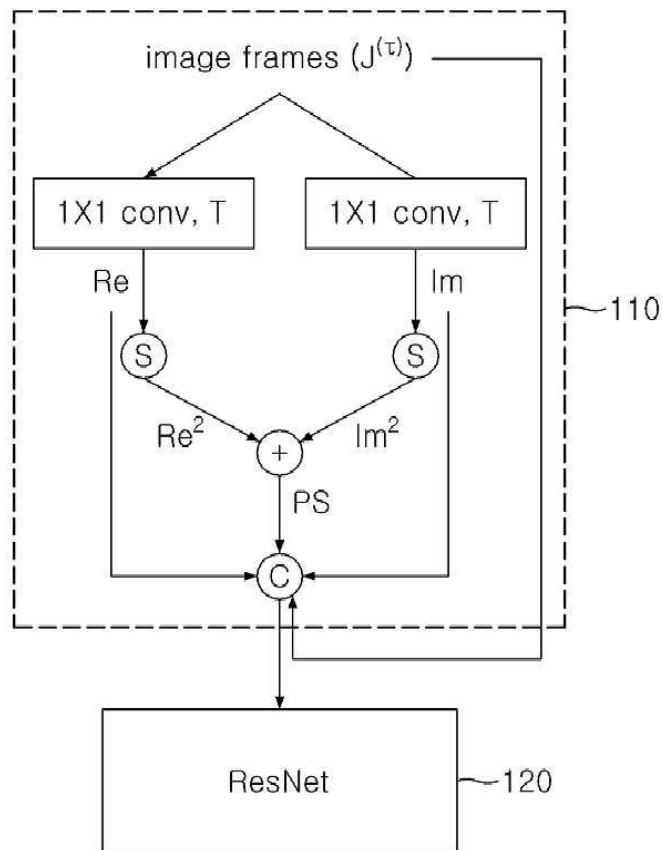
도면2



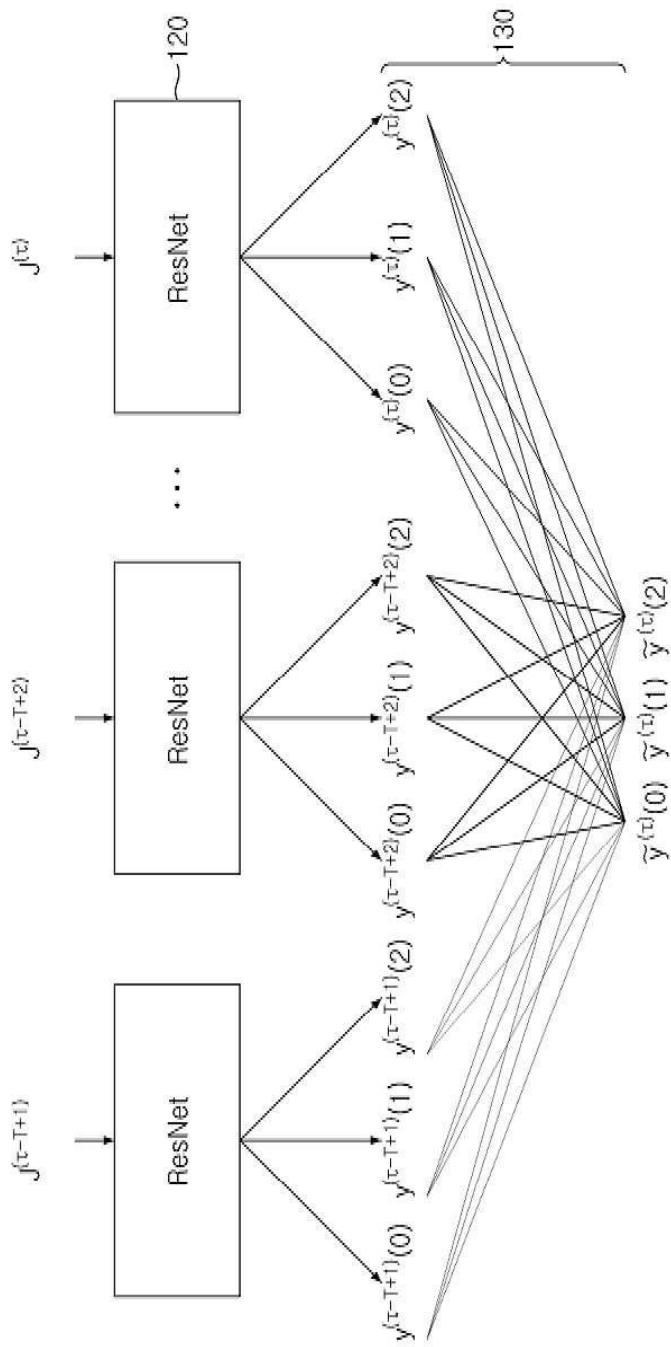
도면3



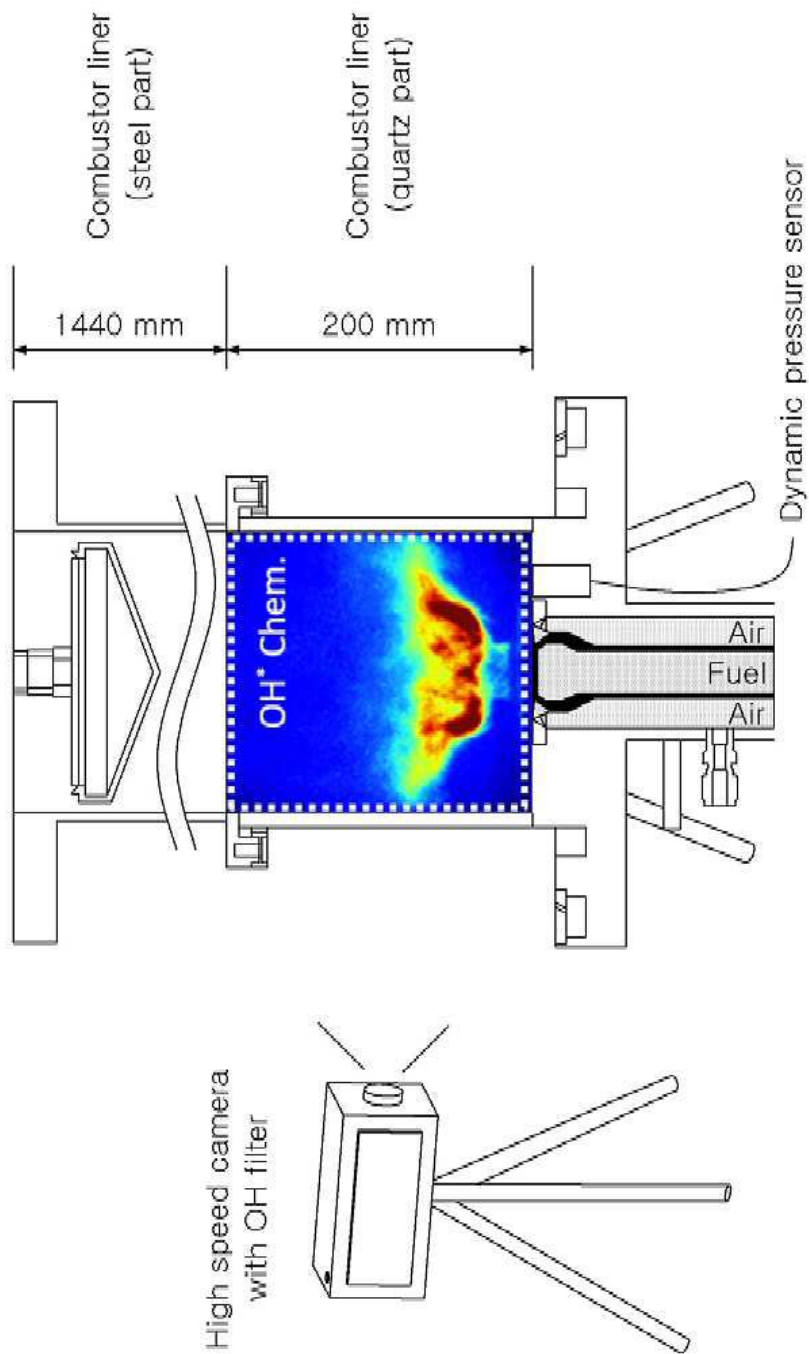
도면4



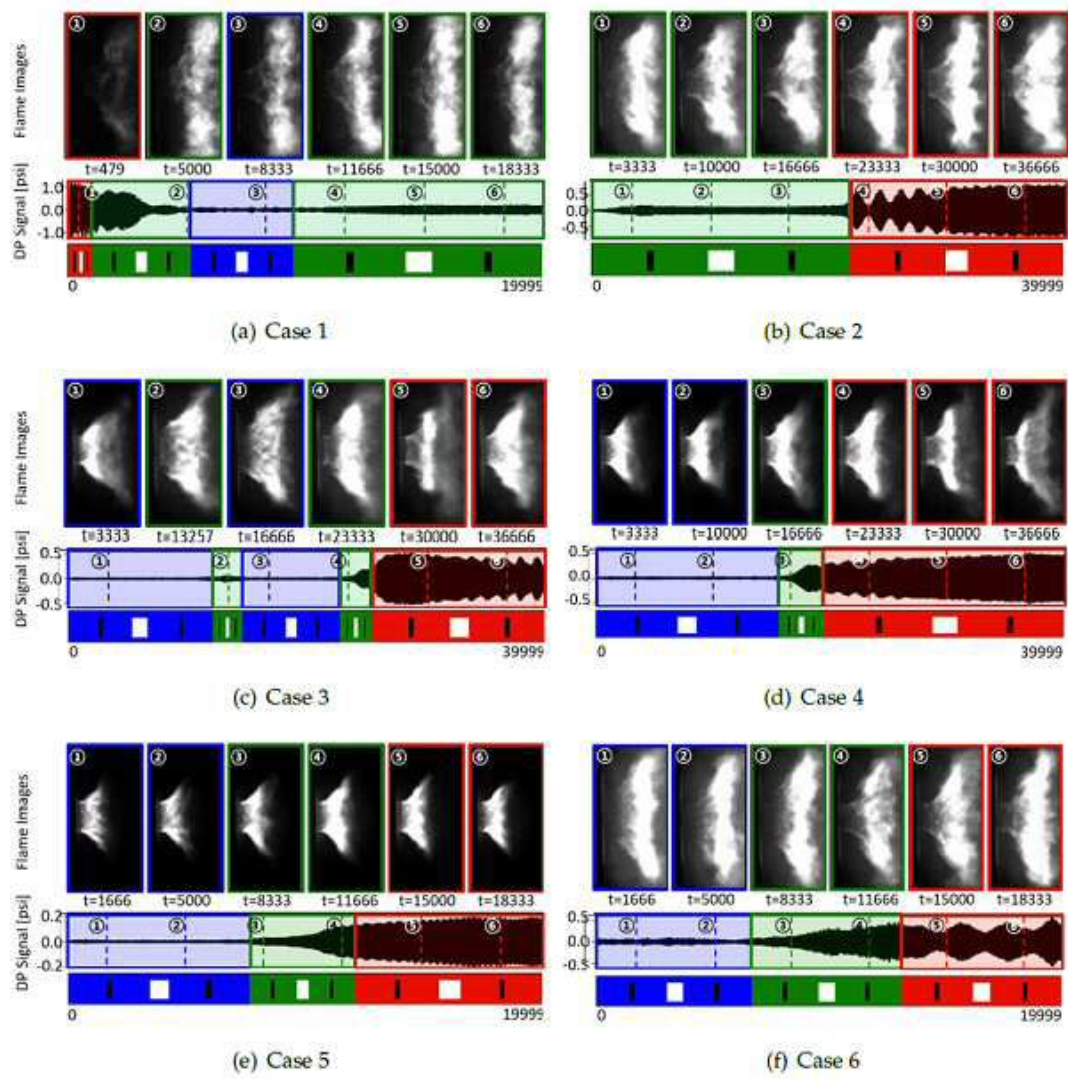
도면5



도면6



도면7



도면8

