



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0023167
(43) 공개일자 2025년02월18일

- | | |
|--|---|
| <p>(51) 국제특허분류(Int. Cl.)
 $G06V\ 10/774$ (2022.01) $G06N\ 3/0464$ (2023.01)
 $G06V\ 10/44$ (2022.01) $G06V\ 10/764$ (2022.01)</p> <p>(52) CPC특허분류
 $G06V\ 10/774$ (2023.08)
 $G06N\ 3/0464$ (2023.01)</p> <p>(21) 출원번호 10-2023-0104287
 (22) 출원일자 2023년08월09일
 심사청구일자 없음</p> | <p>(71) 출원인
 고려대학교 산학협력단
 서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)</p> <p>(72) 발명자
 박유나
 서울특별시 영등포구 문래로 180, 1304호
 윌러번 크리스티안
 02465, 대한민국 서울특별시 동대문구 회기로2길 5, 샤인빌304호</p> <p>(74) 대리인
 김동용</p> |
|--|---|

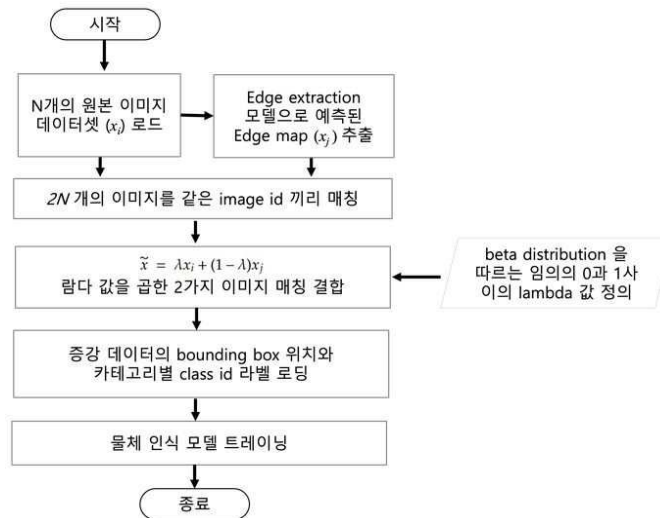
전체 청구항 수 : 총 1 항

(54) 발명의 명칭 **물체 인식 딥러닝 모델의 강인한 학습을 위한 시멘틱 엣지 추출 이미지 기반의 데이터 증강 방법**

(57) 요약

물체 인식 딥러닝 모델의 강인한 학습을 위한 시멘틱 엣지 추출 이미지 기반의 데이터 증강 장치 및 방법이 개시된다. 상기 데이터 증강 방법은 적어도 프로세서(processor)를 포함하는 컴퓨팅 장치에 의해 수행되는 데이터 증강 방법으로서, 업샘플링 기반의 엣지맵 생성 딥러닝 모델에서 추출된 엣지 맵을 추출하여 이미지를 가중 선형 보간(weighted linear interpolation)으로 결합시키는 단계 및 해당 데이터를 이미지 분류 모델로 학습시키는 단계를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06N 3/08 (2023.01)

G06V 10/44 (2023.08)

G06V 10/764 (2023.08)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193973
과제번호	2019-0-00079-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	인공지능대학원지원(고려대학교)
기 여 율	1/1
과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

적어도 프로세서(processor)를 포함하는 컴퓨팅 장치에 의해 수행되는 데이터 증강 방법에 있어서,
 업샘플링 기반의 엣지맵 생성 딥러닝 모델에서 추출된 엣지 맵을 추출하여 이미지를 가중 선형 보간(weighted linear interpolation)으로 결합시키는 단계; 및
 해당 데이터를 이미지 분류 모델로 학습시키는 단계를 포함하는 데이터 증강 방법.

발명의 설명

기술 분야

- [0001] 본 발명은 딥러닝 기반 물체 인식 모델의 강인성을 증대시키고, 물체의 형태 인식에 대한 편향을 증진시키기 위한 엣지 추출 이미지와 원본 이미지를 융합하는 데이터 증강 방법이다. 이 방법은 이미지의 경계를 강화하고 물체의 형태적인 특징을 부각시켜 모델의 성능을 개선하는데 유용하며, 일반화 능력과 객체 인식의 안정성을 향상시키는 데 기여한다.

배경 기술

- [0002] CNN(합성곱 신경망)은 이미지 인식 및 객체 탐지 등의 컴퓨터 비전 작업에 주로 사용되는 딥러닝의 알고리즘으로, 입력 이미지에 3x3, 7x7 과 같은 일정한 사이즈의 작은 필터 (커널)를 이미지 내부에 순차적으로 적용하여 특징을 학습하고 물체의 여부를 인식하거나 물체의 특징을 탐지한다. 딥러닝을 이용한 물체 탐지 및 분류 모델은 보안 및 안전, 자율주행, 의료 진단, 이미지 분석 등 다양한 분야에서 사용되고 있다. 안정적인 적용을 위해 모델의 정확도와 일반화 능력을 개선하는 것이 중요하다.
- [0003] 데이터 증강 기법은 딥러닝에서의 데이터 증강이란 데이터의 양을 늘리기 위해 원본에 각종 변환을 적용하여 딥러닝 모델의 일반화 능력을 향상시키기 위한 주요 방법 중 하나이다. 예를 들어, 학습 데이터셋이 실제 상황에서의 밝기 등의 입력값과 다를 경우 인식을 못하는 경우가 발생할 수 있다. 이 문제는 데이터 증강을 통해서 실제 입력값과 비슷한 데이터 분포를 만들어 내어 실제 상황에서도 정확하게 물체를 인식할 수 있다. 또한, 제한된 데이터 내에서 학습하여 모델을 과적합 시키는 오버피팅 (overfitting) 문제를 줄이는 효과도 있다.
- [0004] 이미지 분류 학습에서 자주 사용되는 데이터 증강 기술로는 MixUp 과 CutMix 가 있다. MixUp 은 두 이미지를 임의의 값을 곱해 linear interpolation 으로 합성하는 증강 방법이며(비특허문헌 1), CutMix 는 이미지의 일부를 잘라내어 다른 이미지에 무작위의 크기와 위치로 섞는 방법이다(비특허문헌 2). 이러한 증강 방법으로 모델의 여러 특성의 이미지 데이터에서 정확도가 더욱 높아지며 일반화 성능이 좋아지는 것이 확인 되었다.
- [0005] 엣지 추출 딥러닝 모델로 Dense Extreme Inception Network 가 선택되었다. 이 모델은 총 6개의 CNN 기반 메인 블록으로 인코더로 구성되어, 각 블록에서 더욱 큰 크기의 샘플을 생성하는 Upsampling 으로 엣지 맵을 생성한다. 해당 모델은 엣지 검출의 평가 기준인 ODS (Object Detection Setting) 와 OIS (Object Interpolation Setting) 의 F-measure 에서 현재까지 가장 높은 정확성의 검출을 보인다. 해당 모델은 사람이 직접 선을 그려 annotate 한 BIPED 데이터셋으로 학습되었다(비특허문헌 3).

선행기술문헌

비특허문헌

- [0006] (비특허문헌 0001) Yun, S., Han, D., Oh, S. J., Chun, S., Choe, J., & Yoo, Y. (2019). CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 6023-6032.
- (비특허문헌 0002) Zhang, Hongyi, et al. "mixup: Beyond empirical risk minimization." arXiv preprint

arXiv:1710.09412 (2017).

(비특허문헌 0003) Poma, Xavier Soria, Edgar Riba, and Angel Sappa. "Dense extreme inception network: Towards a robust cnn model for edge detection." Proceedings of the IEEE/CVF winter conference on applications of computer vision. 2020.

(비특허문헌 0004) Geirhos, Robert, et al. "ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness." arXiv preprint arXiv:1811.12231 (2018).

(비특허문헌 0005) Radford, Alec, et al. "Learning transferable visual models from natural language supervision." International conference on machine learning. PMLR, 2021.

발명의 내용

해결하려는 과제

[0007] 적대적 공격 (adversarial attack)은 딥러닝 모델에 미세한 노이즈를 추가하여 예측 결과가 바뀌는 현상으로, 모델의 신뢰성과 실용성에 영향을 미치는 딥러닝의 문제점 중 하나이다. 이러한 현상은 CNN 알고리즘 특성상 이미지의 픽셀 단위로 필터를 적용하여 local 적인 특징에 의존하여 예측하기 때문에 발생하게 된다. 관련 연구에서는 CNN 모델이 사람과는 달리 물체의 형태 (shape) 보다 이미지의 질감 (texture) 에 따라 이미지를 판단하는 것으로 밝혔다(비특허문헌 4). 따라서 인간이 감지할 수 없는 미세한 노이즈로 이미지의 질감에도 민감한 학습 모델보다 모델의 학습 과정에서 특정 형태에 집중하여 물체의 형태를 잘 인식하고 구분할 수 있도록 형태적인 편향 (shape bias)을 더욱 높여야한다.

[0008] 기존의 데이터 증강 방법들은 이미지를 변환하는 기법을 사용하나, 물체의 형태에 대한 편향 (shape bias) 증진 측면에서 한계가 있다. 현재 가장 보편적으로 쓰이는 CNN 기반의 이미지 분류 모델 ResNet 은 뛰어난 이미지 분류 능력과는 달리 사람이 연필로 그린 그림 (Sketch Dataset) 와 같은 데이터에서는 25.2% 의 정확도를 보인다. 최신 모델 CLIP에서도 스케치 데이터셋에서는 60% 정도의 정확도로 다른 데이터셋에 비해 정확도가 떨어지는 것을 볼 수 있다(비특허문헌 5). 엣지 추출 이미지 학습 방법을 통하여 이미지 분류 모델에 shape bias를 더하여 보다 강건(robust)한 학습 모델을 만들 수 있다.

과제의 해결 수단

[0009] 본 발명의 일 실시예에 따른, 물체 인식 딥러닝 모델의 강인한 학습을 위한 시멘틱 엣지 추출 이미지 기반의 데이터 증강 방법은 적어도 프로세서(processor)를 포함하는 컴퓨팅 장치에 의해 수행되는 데이터 증강 방법으로서, 업샘플링 기반의 엣지맵 생성 딥러닝 모델에서 추출된 엣지 맵을 추출하여 이미지를 가중 선형 보간 (weighted linear interpolation)으로 결합시키는 단계 및 해당 데이터를 이미지 분류 모델로 학습시키는 단계를 포함한다.

발명의 효과

[0010] 본 발명은 적대적 공격에 대응하고 모델의 안정성을 높이기 위해 딥러닝 추출 엣지 맵과 결합하는 데이터 증강 기법을 효과적으로 활용한다. 엣지 특징은 작은 물체나 복잡한 배경을 가진 물체 탐지에도 효과적이다. 또한 추출 모델은 사람이 직접 선을 따서 annotate 된 이미지로 학습되어 추출된 엣지 정보가 추가적으로 학습되기 때문에, 사람이 생각하기에 중요한 물체의 시멘틱 (semantic) 한 정보를 직접적으로 학습할 수 있다. 따라서 본 데이터 증강 방법은 물체 인식 딥러닝 모델의 강인성(robustness)와 형태 인식 (shape bias) 에 대한 편향을 증진시키는데 적합하다. 학습 데이터를 엣지 추출 이미지와 융합하여 다양한 형태의 물체에 대해 강건한 학습을 가능하게 하며, 물체 인식의 안정성과 정확성을 향상시킨다. 도 3을 참조하면, 이미지의 다양한 맥락적 정보의 상황에서 학습 후 성능이 향상된 것을 확인할 수 있었다.

도면의 간단한 설명

[0011] 본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 상세한 설명이 제공된다.

도 1은 본 발명의 일 실시예에 따른 데이터 증강 방법을 설명하기 위한 흐름도이다.

도 2는 데이터 증강 결과 학습 데이터의 예시를 도시한다.

도 3은 딥러닝 물체 인식 모델 RetinaNet의 해당 데이터 증강 학습 효과를 나타내는 그래프이다.

발명을 실시하기 위한 구체적인 내용

- [0012] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시예들에 대해서 특정한 구조적 또는 기능적 설명들은 단지 본 발명의 개념에 따른 실시예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시예들은 다양한 형태로 실시될 수 있으며 본 명세서에 설명된 실시예들에 한정되지 않는다.
- [0013] 본 발명의 개념에 따른 실시예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시예들을 도면에 예시하고 본 명세서에서 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시예들을 특정한 개시 형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물, 또는 대체물을 포함한다.
- [0014] 제1 또는 제2 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 상기 구성 요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소로부터 구별하는 목적으로만, 예컨대 본 발명의 개념에 따른 권리 범위로부터 벗어나지 않은 채, 제1 구성 요소는 제2 구성 요소로 명명될 수 있고 유사하게 제2 구성 요소는 제1 구성 요소로도 명명될 수 있다.
- [0015] 어떤 구성 요소가 다른 구성 요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성 요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성 요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성 요소가 다른 구성 요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는 중간에 다른 구성 요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성 요소들 간의 관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.
- [0016] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 본 명세서에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0017] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0018] 이하, 본 명세서에 첨부된 도면들을 참조하여 본 발명의 실시예들을 상세히 설명한다. 그러나, 특허출원의 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0019] 객체 탐지를 위한 기존 딥러닝 모델을 선정하고, 해당 모델의 pretraining 된 가중치(weight)를 불러온다. 엣지 추출 모델로 입력 이미지에 엣지 맵을 추출한다. 랜덤한 램다 값을 0과 1 사이에서 추출하여 선형 보간에 적용할 가중치를 결정한다. 추출된 이미지의 엣지맵과 원본 이미지를 beta distribution을 따르는 가중 선형 보간(Weighted linear interpolation) 하여 원본 이미지와 추출된 엣지 이미지를 합성하여 증강 데이터를 생성한다. 데이터로더에 물체의 위치(bounding box)와 물체 카테고리 값을 이미지와 함께 제공하여 학습시킨다.
- [0020] 이에 본 발명에서는, 엡샘플링 기반의 엣지맵 생성 딥러닝 모델에서 추출된 엣지 맵을 추출하여 이미지를 가중 선형 보간(weighted linear interpolation)으로 결합시키는 단계, 및 해당 데이터를 이미지 분류 모델로 학습시키는 단계를 포함하는 데이터 증강 방법을 제안한다.
- [0021] 이상에서 설명된 장치는 하드웨어 구성 요소, 소프트웨어 구성 요소, 및/또는 하드웨어 구성 요소 및 소프트웨어 구성 요소의 집합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성 요소는, 예를 들어, 프로세서, 컨트롤러, ALU(Arithmetic Logic Unit), 디지털 신호 프로세서(Digital Signal Processor), 마이크

로컴퓨터, FPA(Field Programmable array), PLU(Programmable Logic Unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(Operation System, OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술 분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(Processing Element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(Parallel Processor)와 같은, 다른 처리 구성(Processing Configuration)도 가능하다.

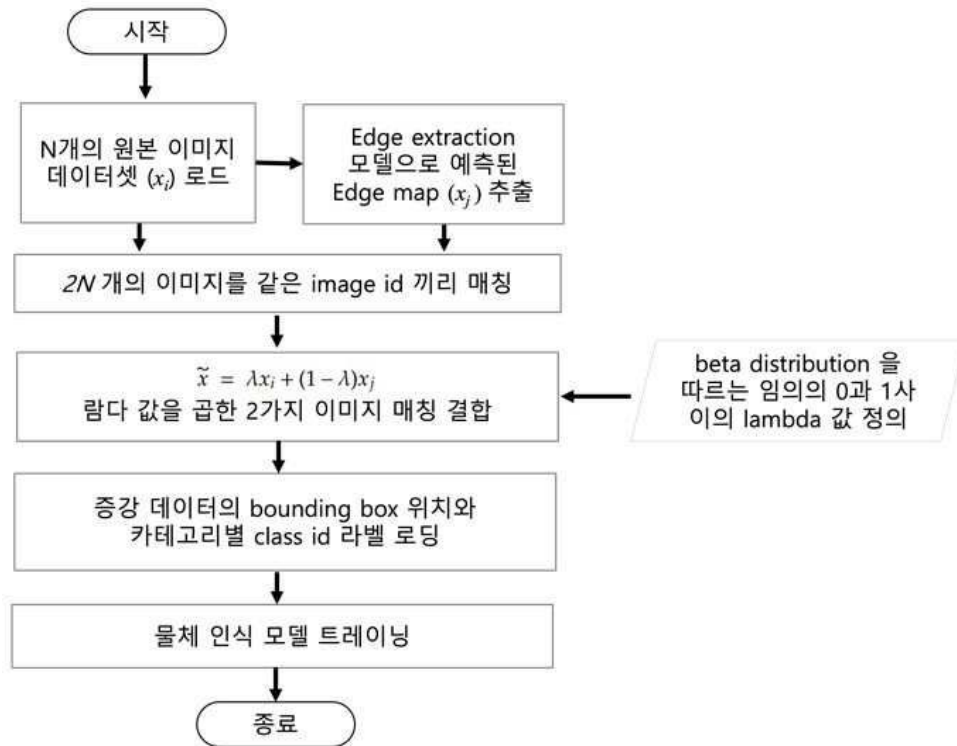
[0022] 소프트웨어는 컴퓨터 프로그램(Computer Program), 코드(Code), 명령(Instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(Collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성 요소(Component), 물리적 장치, 가상 장치(Virtual Equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(Signal Wave)에 영구적으로, 또는 일시적으로 구체화(Embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0023] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(Magnetic Media), CD-ROM, DVD와 같은 광기록 매체(Optical Media), 플롭티컬 디스크(Floptical Disk)와 같은 자기-광 매체(Magneto-optical Media), 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0024] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성 요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성 요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

도면

도면1



도면2



도면3

