



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

第 41 回 碩士學位論文

指導教授 白 俊 基

YOLO v3와 CNN을 사용한 계량기 숫자 검출 및 인식 방법

Meter Digit Detection and Recognition Method
using YOLO v3 and CNN

中央大學校 尖端映像大學院
映像學科 映像工學 디지털이미징 專攻
孫 昌 義
2020年 2月

**YOLO v3와 CNN을 사용한
계량기 숫자 검출 및 인식 방법**
**Meter Digit Detection and Recognition Method
using YOLO v3 and CNN**

이 論文을 碩士學位論文으로 提出함.

2020年 2月

中央大學校 尖端映像大學院
映像學科 映像工學 디지털이미징 專攻
孫 昌 義

孫 昌 義의 碩士學位論文을 認定함

審査委員長 _____ (印)

審査委員 _____ (印)

審査委員 _____ (印)

中央大學校 尖端映像大學院

2020年 2月

목 차

| | |
|---|----|
| 제 1 장 서 론 | 1 |
| 1.1 연구 배경 | 2 |
| 1.2 연구 목적 및 내용 | 5 |
| 1.3 논문 구성 | 6 |
| 제 2 장 배경 이론 | 7 |
| 2.1 1-stage 객체 검출 모델 YOLO | 8 |
| 2.2 인식에 사용되는 대표적 CNN 구조인 VGG | 11 |
| 제 3 장 계량기 숫자 검출 및 인식 방법 | 13 |
| 3.1 계량기 지침/기물 숫자 영역 검출 및 인식 방법의 개요 | 14 |
| 3.2 계량기 지침/기물 영역 검출 | 15 |
| 3.3 숫자 분할 방법 | 17 |

| | |
|--------------------------|-----------|
| 3.4 숫자 인식 방법 | 19 |
| 제 4 장 실험결과 | 20 |
| 4.1 제안하는 알고리즘의 실험결과..... | 21 |
| 제 5 장 결 론 | 30 |
| 5.1 결론 | 31 |
| 참고 문헌..... | 32 |
| 국문 초록 | 36 |
| Abstract..... | 38 |

그 립 목 차

| | |
|--|----|
| 그림 1. 제안하는 방법이 적용되는 환경..... | 4 |
| 그림 2. 2-STAGE 검출기 진행과정..... | 8 |
| 그림 3. 1-STAGE 검출기 진행과정..... | 8 |
| 그림 4. 3x3, 5x5 필터 컨볼루션 사이즈 결과..... | 12 |
| 그림 5. 제안하는 CNN 구조..... | 19 |
| 그림 6. YOLO 지침/기물영역 검출 결과..... | 22 |
| 그림 7. 지침영역에서의 숫자 분할 결과: (A) 입력 이미지, (B) 필터링, 이진화 된 이미지, (C) 연결 성분 라벨링 된 이미지, (D) 최종 진행된 결과 이미지..... | 23 |
| 그림 8. 기물영역에서의 숫자 분할 결과: (A) 입력 이미지, (B) 필터링, 이진화 된 이미지, (C) MSER, 응집 처리된 이미지, (D) 최종 진행된 결과 이미지..... | 24 |
| 그림 9. 제안하는 시스템의 최종 결과 이미지..... | 28 |
| 그림 10. 공용 데이터에 대한 제안하는 시스템의 최종 결과..... | 29 |

표 목 차

| | |
|--|----|
| 표 1. CNN 학습을 위한 지침/기물 숫자 데이터 구성..... | 21 |
| 표 2. YOLO를 사용한 검출 성능..... | 22 |
| 표 3. 제안하는 CNN 구조의 지침 숫자 인식 성능..... | 25 |
| 표 4. VGG16의 지침 숫자 인식 성능..... | 26 |
| 표 5. VGG19의 지침 숫자 인식 성능..... | 26 |
| 표 6. 제안하는 CNN 구조의 기물 숫자 인식 성능..... | 27 |
| 표 7. 숫자 검출 및 인식 시스템 END-TO-END 성능..... | 28 |
| 표 8. 공용 데이터에 대한 숫자 검출 및 인식 성능..... | 29 |

제 1장 서 론

본 장에서는 계량기 검침 자동화 시스템을 위한 숫자 검출 및 인식 방법의 개요와 연구 배경, 목적 및 내용, 그리고 전체적인 논문 구성에 대해 소개한다.

1.1 연구 배경

지난 수십년간 집, 공장, 사무실 등 전국에 가스가 공급되며 사용량을 확인하기 위한 수많은 계량기들이 설치되었다. 계량기에 측정된 수치만큼 소비자가 금액을 지불하는 시스템이기 때문에 검침을 하는 일은 중요하다. 현재 대부분의 검침은 검침원들이 눈으로 보고 숫자를 직접 손으로 입력하는 방식으로 진행되고 있다. 이러한 비효율적인 검침방법을 보다 효율적으로 진행하기 위해서는 검침 자동화 시스템이 필요하다. 이에 대해 전통적인 방법인 이미지 처리 기술부터 최신 방법인 머신러닝, 딥러닝 등 학습 기반의 기술까지 많은 연구가 진행되어 왔다.

가스 검침 시스템을 위하여 다음과 같은 전통적인 방법들이 연구되었다. Anis *et. al.* 은 조명변화(illumination variations)에 민감한 RGB 색공간을 선형 조명(luminance)공간인 YCbCr로 바꾸어 각 파라미터별 범위를 정하고 계량기 검침 영역을 추출한다 [1]. 이러한 방법은 보완을 했을지라도 여전히 빛과 색상에 영향을 많이 받는다는 문제점이 있다. 그 외에도 이미지 처리기반의 방법들인 이진화를 진행한 후 템플릿 매칭(template matching)을 하거나 수평, 수직 방향으로의 픽셀 투영 히스토그램(pixel projections histograms)

을 만들어 검침 영역을 찾는 방법들은 가려짐(occlusion), 회전(rotation) 등에 영향을 받는다는 한계가 있다 [2]-[4].

진행되고 있는 딥러닝 기반의 방법들은 다음과 같다. Yang *et. al.* 은 숫자 분할(digit segmentation) 단계를 사용하지 않고 이미지 전체에 대해 컨볼루션(convolution)을 진행하는 FCN(Fully Convolutional Network)와 어텐션 메커니즘(attention mechanism)을 사용하는 구조를 제안했다 [5]. 이 방법은 충분히 좋은 성능을 갖지만 입력으로 계량기가 포함된 전체 이미지가 아닌 추출된 계량기 검침 영역을 넣어줘야 한다는 한계점이 있다. Laroca *et. al.* 는 네트워크를 간소화한 Fast YOLO 구조와 CR-Net, 다중 작업 학습(multitask learning), CRNN(Convolutional Recurrent Neural Network)을 사용하여 저자들이 제안하는 계량기 데이터의 신빙성을 나타낸다 [6]. 이러한 구조는 본 논문에서 제안하는 시스템의 검침 영역 추출 구조와 유사한 부분이 있다. 하지만 제안하는 방법은 조금 더 나아가 계량기별 갖고 있는 고유값인 기물 영역 추출까지 포함한다.

논문에서 제안하는 방법이 적용되는 환경은 그림 1과 같으며 계량기 검침원이 검침할 숫자를 보고 일일이 입력하는 비효율적인 시스템이다. 이에 본 논문에서는 위의 다양한 연구들과 마찬가지로

기존 시스템 대비 작업이 간소화되어 효율적인 작업이 가능한 자동화 검침 시스템을 만드는 것을 목표로 한다. 비전기반 계량기 검침을 하는데 있어 가장 큰 문제점은 실제 국내에 사용되고 있는 계량기의 종류가 다양하여 지침/기물 영역의 위치와 숫자 모양, 지침 숫자의 수가 달라 인식이 어렵다는 점이다. 때문에 여러 제조회사들이 만든 다양한 계량기 일지라도 강건하게 검출 및 분류가 가능한 딥러닝 기반 알고리즘을 제안한다.

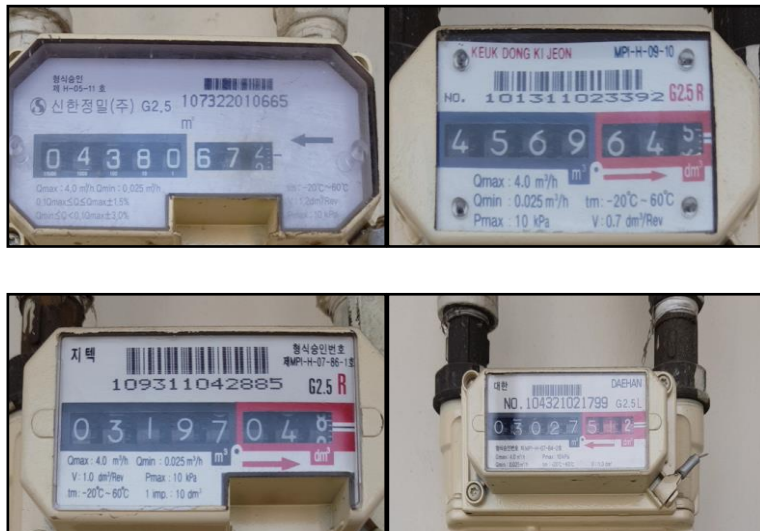


그림 1. 제안하는 방법이 적용되는 환경.

1.2 연구 목적 및 내용

일반적으로 기존의 검침 알고리즘들은 계량기 종류가 다양하여 지침/기물영역의 위치와 숫자 모양, 지침 숫자의 수가 달라 인식이 어렵다. 또한 가려짐, 빛 반사 등의 이유 때문에 성능이 저하된다.

따라서 본 논문에서는 위의 문제들이 있을지라도 강건하게 검출 및 인식이 가능한 딥러닝 기반 end-to-end 알고리즘을 제안한다. 제안하는 방법은 크게 세 부분으로 나눌 수 있다. 먼저 들어온 이미지에서 계량기 지침/기물영역을 사전에 구성한 데이터를 적용시킨 YOLO v3로 추출한다 [7][8]. 다음은 숫자 분할로 지침영역의 경우 연결성분(connected component) 알고리즘 [9], 기물영역의 경우 MSER(Maximally Stable Extremal Regions)기법을 사용한다[10][11]. 마지막으로 분할된 숫자영역에 VGG 기반의 CNN(Convolutional Neural Networks) 네트워크를 구성하여 숫자를 인식한다 [12].

제안하는 방법은 회사별, 버전별로 다양한 타입의 계량기 이미지가 들어와도 강건하게 적용 가능한 딥러닝 기반의 알고리즘이다. 그러므로 검침원들이 기존처럼 검침 및 기물 숫자를 일일이 입력할 필요가 없어 보다 능률적으로 일 할 수 있다.

1.3 논문 구성

본 논문의 구성은 다음과 같다. 2장에서는 객체 검출의 네모 박스 (bounding box)와 클래스 확률(class probability)을 동일한 문제로 두고 동시에 처리하는 대표적인 방법인 YOLO와 인식에 주로 사용되는 대표적인 CNN에 대해 설명하며, 3장에서는 제안하는 계량기 검침/기물 영역의 숫자를 검출 및 인식하는 방법을 설명한다. 4장에서는 실험결과를 제시하고, 마지막으로 5장에서 본 논문의 결론을 맺는다.

제 2장 배경 이론

본 장에서는 제안하는 딥러닝 알고리즘의 기초가 되는 1-stage 객체 검출 알고리즘과 인식의 대표적인 CNN에 대해 설명한다.

2.1 1-stage 객체 검출 모델 YOLO

객체 검출이란 여러 물체에 대해 어떤 물체인지 분류하는 인식 문제와 그 물체가 어디 있는지 네모 박스를 통해 위치 정보를 나타내는 localization 문제를 둘 다 해내야 하는 분야이다. 딥러닝을 이용한 객체 검출은 크게 1-stage 검출기와 2-stage 검출기로 나눌 수 있다. 2-stage 검출기의 경우 그림 2처럼 위치 제안(regional proposal)과 인식이 순차적으로 이루어지고, 1-stage 검출기는 그림 3처럼 두 단계가 동시에 이루어진다.

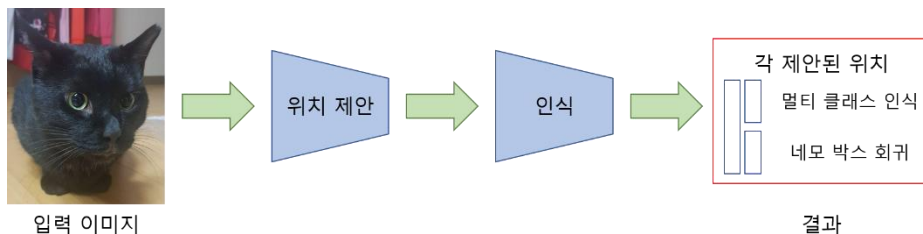


그림 2. 2-stage 검출기 진행과정.

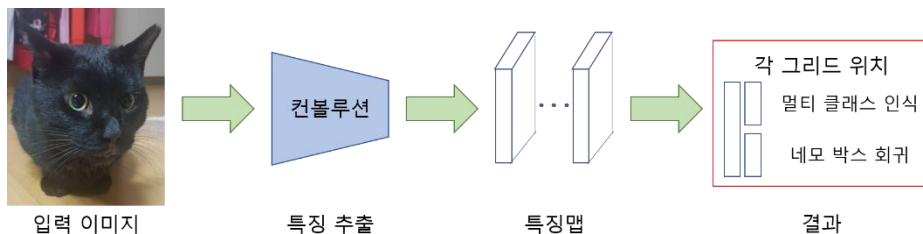


그림 3. 1-stage 검출기 진행과정.

You Only Look Once(YOLO)는 1-stage 구조의 대표적 알고리즘으로 하나의 컨볼루션 네트워크(single convolutional network)를 통해 여러 네모 박스(multiple bounding box)에 대한 클래스 확률을 계산하는 방식이다.

우선 입력 이미지를 $S \times S$ 그리드(grid)로 나누고, 각각의 그리드 셀(grid cell)은 B개의 네모박스와 각 박스에 대한 신뢰도 점수(confidence score)를 갖는다.

$$\text{ConfidenceScore} = p(obj) * IoU'_p, \quad (1)$$

각각의 그리드 셀은 C개의 조건부 클래스 확률(conditional class probability)를 갖는다.

$$\text{ConfidenceClassProbability} = p(class / obj), \quad (2)$$

각 네모박스는 박스의 중심으로 그리드 셀의 범위에 대한 상대값인 (x,y) 좌표와 전체 이미지의 폭, 높이에 대한 상대값인 (w,h) 와 신뢰도로 구성된다.

테스트 시간에는 조건부 클래스 확률과 네모박스의 신뢰도 점수를 곱하여 정확한 클래스 신뢰도 점수(class-specific confidence score)를 얻는다.

$$\begin{aligned}\text{ClassSpecificConfidenceScore} &= p(\text{class} / \text{obj}) * p(\text{obj}) * IoU_p^t \\ &= p(\text{class}) * IoU_p^t, \quad (3)\end{aligned}$$

2.2 인식에 사용되는 대표적 CNN구조인 VGG

VGGNet은 옥스포드 대학의 연구팀에서 개발된 모델로 2014년 이미지넷 이미지 인식 대회에서 준우승을 한 모델이다. 여기서 말하는 VGGNet은 16개 또는 19개의 층으로 구성된 모델인 VGG16, VGG19를 의미한다.

VGGNet의 핵심은 네트워크의 깊이를 깊게 만드는 것이 성능에 어떤 영향을 미치는지 확인하고자 한 것이다. 깊이의 영향만을 최대한 확인하고자 컨볼루션 필터커널의 사이즈는 가장 작은 3×3 으로 고정했다. 필터커널의 사이즈가 크면 이미지의 사이즈가 빠르게 축소되기 때문에 네트워크 깊이를 깊게 만들기 불가능하기 때문이다. 또한 그림 4처럼 3×3 필터로 두 차례 컨볼루션을 하는 것과 5×5 필터로 한번 컨볼루션 하는 것이 대응되지만 필요 가중치의 수 ($3 \times 3 \times 2 = 18, 5 \times 5 \times 1 = 25$)가 더 적어 학습 속도가 빨라진다. 동시에 층의 개수가 늘어나면서 특성에 비선형성을 더 증가시키기 때문에 특성이 점점 더 유용해진다.

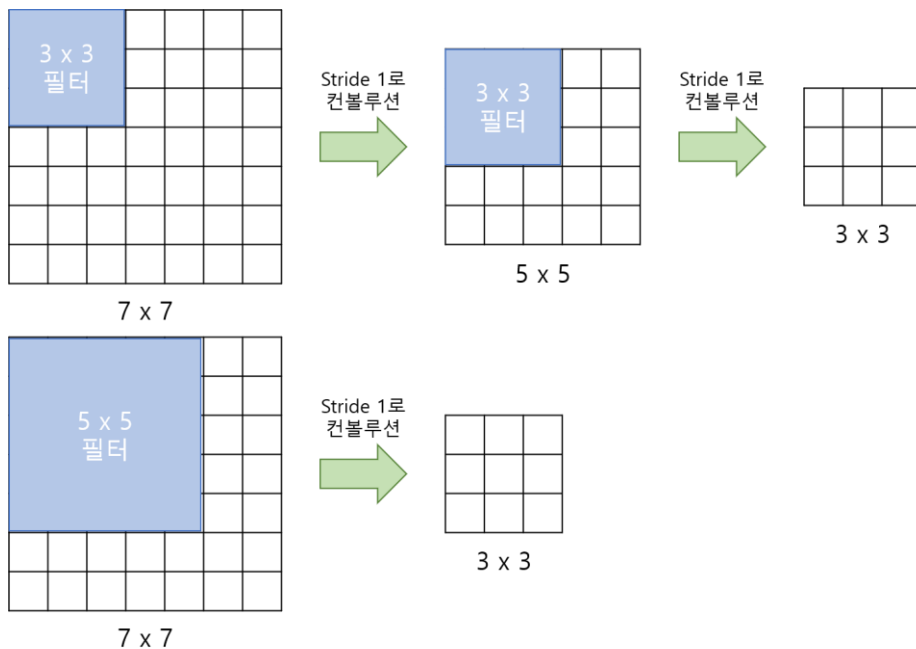


그림 4. 3x3, 5x5 필터 컨볼루션 사이즈 결과.

VGG 네트워크는 입력 이미지를 3x3 필터커널로 컨볼루션하는데 제로 패딩(zero padding)은 1만큼 해주고, 컨볼루션 보폭(stride)는 1로 설정해준다. 활성화시키기 위해 ReLU 함수가 적용되며 마지막 층을 제외하고는 항상 적용된다. 다음 단계의 컨볼루션에서는 2x2 최대 풀링(max pooling)을 보폭 2로 적용해 진행한다. 위의 두 단계가 반복된 후 fully connected layer 단계를 거친다. 마지막 층은 클래스의 개수인 1000개의 뉴런으로 구성하며 softmax 함수로 활성화시킴으로써 1000개의 클래스를 분류할 수 있다.

제 3장 계량기 숫자 검출 및 인식 방법

본 장에서는 제안하는 계량기 검침 자동화 시스템을 위한 딥러닝 기반 end-to-end 방식의 계량기 지침/기물 영역 검출, 숫자 분할 및 숫자 인식 방법에 대해 설명한다.

3.1 계량기 지침/기물 숫자 영역 검출 및 인식 방법의 개요

실제 전국에 설치되어 있는 계량기들은 지침/기물 영역 위치와 숫자 수가 회사별, 버전별로 다르다는 특징이 존재한다. 때문에 입력 이미지로부터 최종 숫자 인식까지 한번에 진행하는 알고리즘 구성으로 도출하기에는 어려움이 있다. 본 논문에서는 계량기 이미지가 들어왔을 때 지침/기물 영역 검출과 숫자 분할 및 숫자 인식 과정인 3단계의 구조로 진행한다.

3.2 계량기 지침/기물 영역 검출

계량기의 종류마다 지침/기물 영역 위치 및 크기는 다르다. 계량기의 지침/기물 영역을 추출하기 위해서는 이처럼 여러 타입에 강건한 방법이 필요하다. 딥러닝은 여러 비선형 변환기법(nonlinear transformation or activation)의 조합을 통해 높은 수준의 추상화를 시도하는 알고리즘의 집합이며, 큰 틀에서 사람의 사고방식을 컴퓨터에게 가르치는 한 분야라고 이야기할 수 있다. 이러한 딥러닝 구조를 사용하면 가려짐, 회전, 유사한 모양에 강건하게 객체 검출이 가능하다.

객체 검출은 물체의 위치를 찾는 localization과 물체를 식별하는 인식을 합하여 진행한다. 이러한 두 과정을 동시에 진행함으로써 처리 속도는 빠르지만 정확도가 비교적 낮은 방법인 1-stage 구조 검출기 YOLO, SSD[13] 계열이 있다. 두 과정을 순차적으로 진행하며 처리 속도는 느리지만 정확도는 비교적 높은 방법인 2-stage 구조 검출기 R-CNN[14]-[16] 계열이 있다. 계량기 검침을 정확하게 하기 위해 검침원이 촬영을 한 후 최종 결과를 확인하는 과정이 필요하기 때문에 빠른 처리속도가 필수적이다. 그러므로 처리속도가 보다 빠른 1-stage 구조를 사용하며 이중에서도 가장 빠른 YOLO를 사용한다. 하지만 YOLO v1은 작은 객체에 약하다는 단점이 존재하며

계량기의 지침/기물 영역은 작은 편에 속한다. 이런 단점을 가장 잘 보완하기 위해 멀티 스케일 예측(multi-scale prediction)을 추가하여 성능을 높인 YOLO v3를 본 논문에서는 사용한다.

3.3 숫자 분할 방법

앞의 YOLO v3 detector를 통해 지침/기물 영역을 검출했다. 검출한 지침영역내의 숫자들은 어느 정도의 간격을 두고 존재하며 각 숫자는 하나의 폐영역을 갖고 있다는 공통적 특징이 있다. 반면 계량기 타입에 따라 지침해야 할 숫자가 4~6으로 정확히 알 수 없다는 어려움도 존재한다. 따라서 숫자의 수에 관계없이 각 폐영역을 갖는 숫자에 대한 영역을 나누는 방법이 필요하다. 제안하는 시스템에서 객체를 분별하기 위해 인접한 픽셀 값들끼리 그룹화 하여 번호를 매기는 라벨링 방법 중 하나인 연결 성분 알고리즘을 사용한다. 주로 컴퓨터 비전환경에서 사용되는 연결 성분 라벨링 (connected component labeling) 방법은 이진화 디지털 이미지에서 연결된 영역들을 검출하는데 사용되며 검출해야 하는 숫자의 개수에 크게 영향을 받지 않는다. 그러므로 필터링을 통한 영상 개선을 진행 후 연결 성분 라벨링을 통해 지침영역 내의 숫자를 검출을 진행한다.

기물영역의 숫자 간 간격이 매우 좁으며 바로 위의 바코드와 거의 인접해 있는 경우가 있기 때문에 지침영역에서 사용한 연결 성분 알고리즘을 사용했을 경우 인접한 영역과 동일한 라벨이 부여될 수 있다. 또한 지침영역의 숫자보다 크기가 작아 빛과 같은 주변의

영향에 보다 취약하다는 어려움이 있다. 이러한 어려움 때문에 지침 영역에서 사용한 연결 성분 알고리즘을 사용할 수 없다. 제안하는 시스템에서는 이처럼 숫자의 크기가 작고, 인접하게 타 영역이 존재하며 빛의 영향을 받는 상황에서 보다 강건하게 영역을 검출할 수 있는 방법인 MSER을 사용한다. MSER은 입력 영상에서 주변의 영역들과 비교하여 밝기, 색과 같은 속성들이 다른 영역을 검출함으로써 기물영역 내의 숫자를 검출하는데 사용하며 바코드와 같은 영역은 필터링을 통해 제거한다.

3.4 숫자 인식 방법

제안하는 시스템의 마지막 단계로 앞 단계까지의 진행을 통해 얻은 지침/기물 영역의 숫자 영역들에 인식과정이다. 숫자 인식에 대한 방법으로는 마르코프 모델(Markov model)을 사용하는 방법에서부터 CNN인 딥러닝을 사용하는 방법까지 많은 기법들이 있다 [17][18]. 제안하는 시스템에서는 많은 숫자 인식 방법 중 인식을 하기 위한 대표적인 딥러닝 기반 VGG 네트워크 구조를 변형하여 사용한다. 기존 VGG 모델처럼 3×3 컨볼루션 2번 또는 3번씩 진행하면 지침/기물 영역 내의 숫자 영역 크기가 작기 때문에 최종 특징맵이 너무 작아지는 결과가 나오게 된다. 따라서 그림 5에서 보듯이 제안하는 네트워크는 VGG처럼 3×3 컨볼루션을 2번씩 진행하되 전체적인 깊이(depth)를 줄였다. 이처럼 보다 가볍게 만든 VGG기반 CNN과정을 통해 인식을 진행한다.



그림 5. 제안하는 CNN 구조.

제 4장 실험결과

본 장에서는 3장에서 제안한 계량기 지침/기
물 숫자 검출 및 인식 방법의 실험결과를 확인
한다.

4.1 제안하는 알고리즘의 실험결과

이번 장에서는 제안된 계량기 검침 시스템에 대한 실험적 성능을 보여준다. 실험에 사용된 학습 및 테스트는 다양한 타입의 계량기 촬영을 통해 얻었으며 320x440 크기로 구성된 데이터와 2019년 2월에 전자 저널(electronics journal)에서 출판된 Laroca *et. al.* 논문에서 구성한 UFPR AMR 공용 데이터로 진행했다 [19]. 지침, 기물 영역 검출을 위한 학습 데이터는 5000개를 각 영역별 레이블링 작업을 진행하여 구성했다. 숫자 인식을 위한 학습 데이터는 검출된 지침영역내 숫자 14,306개, 기물영역내 숫자 23,759개로 표 1처럼 구성했다. 학습 및 테스트에 사용된 컴퓨터 GPU 사양은 GTX 1080Ti이다.

표 1. CNN 학습을 위한 지침/기물 숫자 데이터 구성.

| | 0 | 1 | 2 | 3 | 4 |
|----|-------|-------|-------|-------|-------|
| 지침 | 1,880 | 1,921 | 1,765 | 1,632 | 1,812 |
| 기물 | 2,270 | 2,542 | 2,491 | 2,994 | 2,352 |
| | 5 | 6 | 7 | 8 | 9 |
| 지침 | 1,924 | 1,729 | 1,643 | 1,540 | 1,560 |
| 기물 | 2,108 | 2,453 | 2,243 | 2,010 | 2,296 |

첫번째 실험으로, 제안하는 검침 시스템의 지침 및 기물 영역 검출 성능에 대해 테스트했다. 5,000개의 이미지로 학습된 YOLO v3의 결과는 그림 6, 성능은 표 2와 같다. 여기서 참(true)의 기준은 지침 영역을 포함하며 다른 문자가 들어오지 않는 범위이다.

표 2. YOLO를 사용한 검출 성능.

| | 지침 영역 | 기물 영역 | 지침+기물영역 |
|----|---------|---------|---------|
| 참 | 262 | 243 | 228 |
| 전체 | 266 | 266 | 266 |
| 성능 | 98.496% | 91.353% | 85.714% |

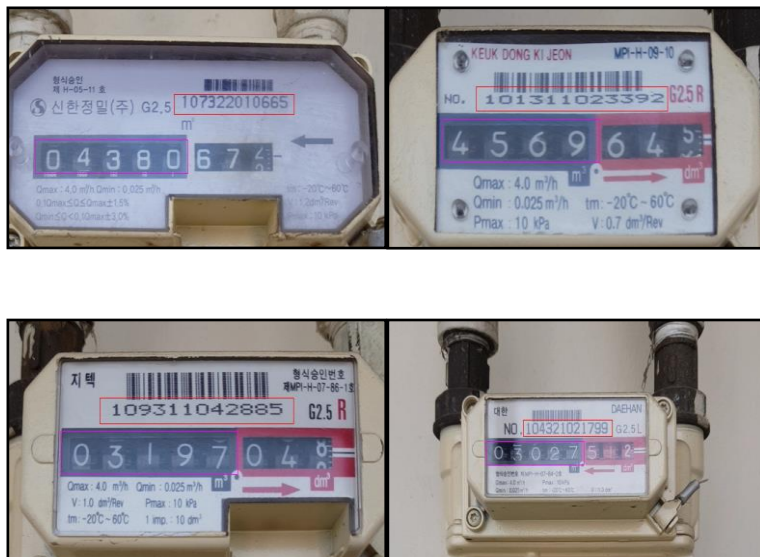


그림 6. YOLO 지침/기물영역 검출 결과.

다음으로 그림 7은 이어지는 내용과 같은 지침영역의 숫자 분할 진행결과이다. 우선 지침 영역이 입력으로 들어와 양방향(bilateral), 모폴로지(morphology)와 같은 필터링 과정을 거쳐 이진화를 진행한다. 이후 연결 성분 라벨링을 사용 후 예외처리와 검출된 박스들의 중간사이즈를 기준(refer)으로 두고 크기 재조정(rescale)과정을 거쳐 최종 숫자 영역을 검출한다.

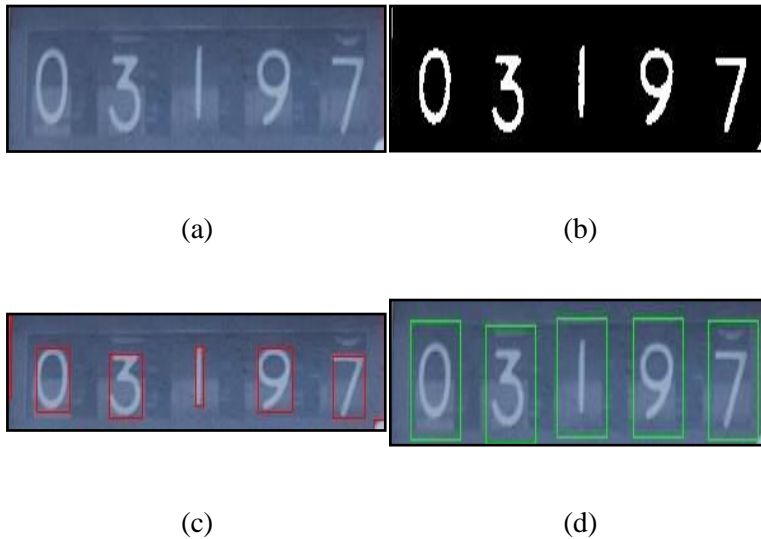


그림 7. 지침영역에서의 숫자 분할 결과: (a) 입력 이미지, (b) 필터링, 이진화 된 이미지, (c) 연결 성분 라벨링 된 이미지, (d) 최종 진행된 결과 이미지.

기물영역의 숫자 분할 과정은 이어지는 내용과 같으며 그림 8은 진행결과이다. 기물 영역의 경우 지침 영역과 색이 반대 이므로 색 반전 과정이 추가되며 이진화는 진행하지 않는다. 이후 MSER을 사용해 특징이 있는 박스를 검출한 후 예외처리 한다. 추가로 높이(height) 기준 스무딩된 히스토그램(smoothed histogram)의 최대값을 기준으로 잡아 예외처리 하며 폭(width)도 마찬가지로 진행한다. 최종 박스는 여러 개가 중복된 형태인데 중복된 영역을 하나로 통합해주는 NMS(Non Maximum Suppression)을 사용해 최종 숫자 영역을 검출한다.

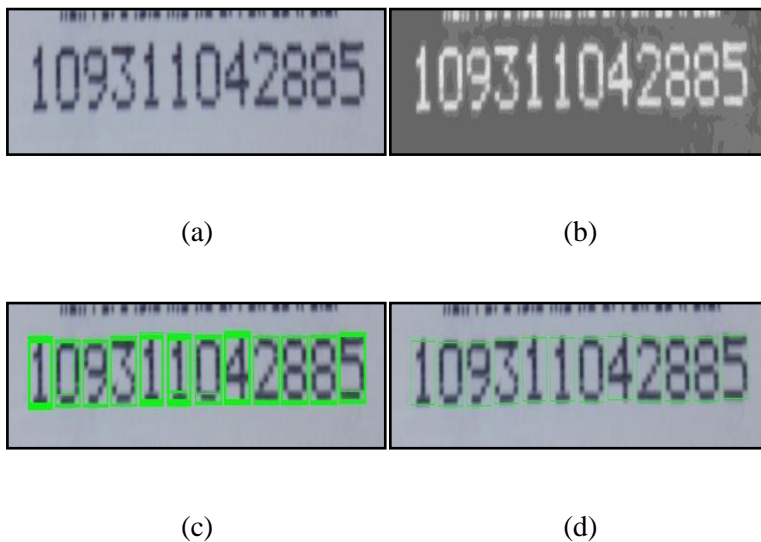


그림 8. 기물영역에서의 숫자 분할 결과: (a) 입력 이미지, (b) 필터링, 이진화 된 이미지, (c) MSER, 응접 처리된 이미지, (d) 최종 진행된 결과 이미지.

마지막 실험으로 제안하는 분류 네트워크인 CNN 성능이다. CNN의 인식 성능을 평가하기 위해 학습 데이터와 다른 527장의 이미지의 지침 및 기물영역 내의 숫자들을 잘라(crop) 진행했다. 표 3은 지침 영역의 숫자들에 대한 인식 결과이며 총 숫자에 대한 성능은 99.59%이다. 기존에 인식으로 많이 쓰이는 VGGNet에서 표 4는 VGG16 결과이며 전체적 성능은 97.37%, 표 5는 VGG19 결과이며 전체적 성능은 97.98%이다. 제안한 구조, VGG16, VGG19에 대한 필요 파라미터는 각각 323만, 1735만, 2266만으로 제안한 네트워크가 보다 더 효율적임을 알 수 있다.

표 3. 제안하는 CNN 구조의 지침 숫자 인식 성능(99.59%).

| | 0 | 1 | 2 | 3 | 4 |
|----|--------|--------|--------|--------|------|
| 참 | 1169 | 430 | 488 | 383 | 284 |
| 전체 | 1175 | 432 | 492 | 384 | 284 |
| 성능 | 99.49% | 99.54% | 99.19% | 99.74% | 100% |
| | 5 | 6 | 7 | 8 | 9 |
| 참 | 270 | 230 | 234 | 215 | 193 |
| 전체 | 271 | 231 | 235 | 215 | 193 |
| 성능 | 99.63% | 99.56% | 99.57% | 100% | 100% |

표 4. VGG16의 지침 숫자 인식 성능(97.37%).

| | 0 | 1 | 2 | 3 | 4 |
|----|--------|--------|--------|--------|--------|
| 참 | 1149 | 430 | 485 | 371 | 280 |
| 전체 | 1175 | 432 | 492 | 384 | 284 |
| 성능 | 97.79% | 99.54% | 98.58% | 96.61% | 98.59% |
| | 5 | 6 | 7 | 8 | 9 |
| 참 | 247 | 221 | 235 | 207 | 186 |
| 전체 | 271 | 231 | 235 | 215 | 193 |
| 성능 | 91.14% | 95.67% | 100% | 96.28% | 96.37% |

표 5. VGG19의 지침 숫자 인식 성능(97.98%).

| | 0 | 1 | 2 | 3 | 4 |
|----|--------|--------|--------|--------|--------|
| 참 | 1155 | 431 | 488 | 364 | 280 |
| 전체 | 1175 | 432 | 492 | 384 | 284 |
| 성능 | 98.30% | 99.77% | 99.19% | 94.79% | 98.59% |
| | 5 | 6 | 7 | 8 | 9 |
| 참 | 264 | 223 | 234 | 208 | 186 |
| 전체 | 271 | 231 | 235 | 215 | 193 |
| 성능 | 97.42% | 96.54% | 99.57% | 96.74% | 96.37% |

표 6은 제안하는 네트워크를 사용한 기물 영역 숫자들의 인식 결과이며 총 숫자에 대한 성능은 97.7%이다.

표 6. 제안하는 CNN 구조의 기물 숫자 인식 성능.

| | 0 | 1 | 2 | 3 | 4 |
|----|--------|--------|--------|--------|------|
| 참 | 1712 | 2489 | 2407 | 1104 | 499 |
| 전체 | 1722 | 2494 | 2408 | 1107 | 499 |
| 성능 | 99.42% | 99.80% | 99.96% | 99.73% | 100% |
| | 5 | 6 | 7 | 8 | 9 |
| 참 | 465 | 462 | 377 | 601 | 1241 |
| 전체 | 466 | 466 | 381 | 602 | 1241 |
| 성능 | 99.79% | 99.14% | 98.95% | 99.83% | 100% |

표 7은 지침 및 기물영역 검출에서부터 숫자 인식까지의 end-to-end 성능을 보여주며 그림 9는 제안하는 시스템의 최종 결과 이미지이다. End-to-end 성능을 측정할 때 지침, 기물 영역 내의 숫자 모두 정확하게 검출 및 인식하는 경우에만 참이라고 평가했다.

표 7. 숫자 검출 및 인식 시스템 end-to-end 성능.

| | 지참 | 기물 | 지참+기물 |
|----|--------|--------|--------|
| 참 | 238 | 202 | 182 |
| 전체 | 266 | 266 | 266 |
| 성능 | 89.47% | 75.94% | 68.42% |



지참 [0,4,3,8,0],
기물: [1,0,7,3,2,2,0,1,0,6,6,5]

지참 [4,5,6,9],
기물: [1,0,1,3,1,1,0,2,3,9,2]



지참 [0,3,1,9,7],
기물: [1,0,9,3,1,1,0,4,2,8,8,5]

지참 [0,3,0,2,7],
기물: [1,0,4,3,2,1,0,2,1,7,9,9]

그림 9. 제안하는 시스템의 최종 결과 이미지.

다음으로 공용데이터인 UFPR AMR로의 최종 성능 결과이다. 데이터 구성은 학습데이터 800장, 검증데이터 400장, 테스트데이터 800장이다. 표 8은 지침영역에 대한 end-to-end 성능이며, 그림 10 최종 결과 이미지이다.

표 8. 공용 데이터에 대한 숫자 검출 및 인식 성능.

| | 지침 |
|----|--------|
| 참 | 641 |
| 전체 | 800 |
| 성능 | 80.13% |



지침: [1,3,1,2,8]

지침: [2,3,5,5,5]

그림 10. 공용 데이터에 대한 제안하는 시스템의 최종 결과.

제 5장 결 론

본 장에서는 제안한 계량기 지침/기물 영역
숫자 검출 및 인식 방법의 장단점을 간략히 요
약하고, 향후 진행될 과제에 대해 기술한다.

5.1 결론

본 논문에서는 딥러닝 기반의 기술을 기반으로 효율적인 검침을 위한 자동화 검침 시스템을 나타낸다. 컴퓨터 비전 기반으로 계량기를 영역 검출 및 인식을 수행하기에는 계량기 종류별 각기 다른 지침 및 기물 영역과 다른 숫자 모양을 갖고 있다는 어려움이 존재한다. 본 시스템은 딥러닝 기반 방법을 사용하여 어려움을 해결했으며 3단계 구조의 지침/기물 영역 검출, 숫자 분할, 인식 단계로 진행된다. 우선 YOLO 영역 검출기를 사용함으로써 각기 다른 지침, 기물 영역을 검출했다. 다음으로 검출된 지침, 기물 영역에 대해 숫자간 약간의 간격이 존재, 숫자가 보다 작아 주변의 영향을 많이 받는다는 특징에 맞게 연결 성분 라벨링과 MSER을 사용했다. 마지막으로 VGG 기반의 CNN 구조를 만들어 인식을 진행했다.

위에서 언급했듯이 검침원이 일일이 검침 숫자를 입력하는 기존의 방법은 많은 인력이 요구되며 비효율적이다. 제안하는 시스템은 입력으로 계량기가 포함된 이미지가 들어오면 이를 컴퓨터 비전, 딥러닝 기반의 방법을 사용하여 지침/기물 숫자를 인식하는 효율적인 검침 방법을 제공할 수 있다.

참고 문헌

- [1] Atif Anis, Md. Khaliluzzaman, Mohammad Yakub, Niloy Chakraborty, Kaushik Deb, “Digital electric meter reading recognition based on horizontal and vertical binary pattern”, 3rd IEEE International Conference on Electrical Information and Communication Technology (EICT), December 2017.
- [2] Lamiaa A. Elrefaei, Asrar Bajaber, Sumayyah Natheir, Nada AbuSanab, Marwa Bazi, “Automatic electricity meter reading based on image processing”, IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), November 2015.
- [3] Dongmei Shu, Shuhua Ma, Chunguo Jing, “Study of the Automatic Reading of Watt Meter Based on Image Processing Technology”, 2nd IEEE Conference on Industrial Electronics and Applications, May 2007.
- [4] Hitesh Rajput, Tanmoy Som, Soumitra Kar, “An Automated Vehicle License Plate Recognition System”, Computer, Vol. 48, Issue 8, pp. 56-61, August 2015.
- [5] Fan Yang, Lianwen Jin, Songxuan Lai, Xue Gao, Zhaohai Li, “Fully Convolutional Sequence Recognition Network for Water Meter Number Reading”, IEEE Access, Vol. 7, pp. 11679-11687, January 2019.

- [6] Rayson Laroca, Victor Barroso, Matheus A. Diniz, Gabriel R. Gonçalves, William R. Schwartz, David Menotti, “Convolutional neural networks for automatic meter reading”, *Journal of Electronic Imaging*, Vol. 28, January 2019.
- [7] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, “You Only Look Once: Unified, Real-Time Object Detection”, *IEEE Conference on Computer Vision and Pattern Recognition*, May 2016.
- [8] Joseph Redmon, Ali Farhadi, “YOLOv3: An Incremental Improvement”, *IEEE Conference on Computer Vision and Pattern Recognition*, April 2018.
- [9] Costantino Grana, Daniele Borghesani, Rita Cucchiara, “Connected Component Labeling Techniques on Modern Architectures”, *International Conference on Image Analysis and Processing*, pp. 816-824, 2009 Springer.
- [10] Ron Kimmel, Cuiping Zhang, Alex Bronstein, Michael Bronstein, “Are MSER Features Really Interesting?”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 33, Issue. 11, pp. 2316-2320, June 2016.
- [11] Huizhong Chen, Sam S. Tsai, Georg Schroth, David M. Chen, Radek Grzeszczuk, Bernd Girod, “Robust text detection in natural images with edge-enhanced Maximally Stable Extremal Regions”, *18th IEEE International Conference on Image Processing*, September 2011.
- [12] Karen Simonyan, Andrew Zisserman, “Very Deep Convolutional

Networks for Large-Scale Image Recognition”, IEEE Conference on Computer Vision and Pattern Recognition, April 2015.

- [13] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg, “SSD: Single Shot MultiBox Detector”, European Conference on Computer Vision, pp. 21-37, September 2016.
- [14] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, “Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation”, IEEE Conference on Computer Vision and Pattern Recognition, pp. 580-587, October 2014.
- [15] Ross Girshick, “Fast R-CNN”, The IEEE International Conference on Computer Vision (ICCV), pp. 1440-1448, 2015.
- [16] Shaoging Ren, Kaiming He, Ross Girshick, Jian Sun, “Faster R-CNN: Toward Real-Time Object Detection with Region Proposal Networks”, Advances in Neural Information Processing Systems 28 (NIPS), 2015.
- [17] Tran Duc Duan, Tran Le Hong Du, Trần Vĩnh Phước, Nguyen Viet Hoang, “Building an Automatic Vehicle License-Plate Recognition System”, International Conference in Computer Science, February 2005.
- [18] Ian J. Goodfellow, Yaroslav Bulatov, Julian Ibarz, Sacha Arnoud, Vinay Shet, “Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks”, IEEE

Conference on Computer Vision and Pattern Recognition, April 2014.

- [19] Rayson Laroca, Victor Barroso, Matheus A. Diniz, Gabriel R. Gonçalves, William Robson Schwartz, David Menotti, “Convolutional Neural Networks for Automatic Meter Reading”, IEEE Conference on Computer Vision and Pattern Recognition, February 2019.

국문 초록

YOLO v3와 CNN을 사용한 계량기 숫자 검출 및 인식 방법

중앙대학교 첨단영상대학원
영상학과 영상공학 디지털이미징 전공
손 창 의

계량기 검침 분야는 이미지 처리 기술 중심인 전통적인 방법부터 머신러닝, 딥러닝과 같은 학습 기반의 기술을 사용하는 최신 방법까지 많은 연구들이 진행되어 왔다. 컴퓨터 비전기반 계량기 검침의 가장 큰 문제점은 다양한 계량기의 종류에 대해 인식이 어렵다는 점이다. 실제로 국내에 설치 되어있는 계량기의 주된 제조회사만해도 5곳 이상이며 버전별로 계량기 지침, 기물 영역의 위치 및 숫자 모양이 다르다. 이러한 문제 때문에 여전히 대부분의 검침은 검침원들이 눈으로 보고 숫자를 직접 손으로 입력하는 방식으로 진행되고 있다. 본 논문에서는 사람이 일일이 확인해야하는 기존의 검침 시스템 대비 간소화되고 효율적인 작업이 가능한 자동화 계량기 검침 방법을 제안한다. 제안하는 방법은 크게 3부분으로 구성되어 있다: i) YOLO v3를 사용한 지침 및 기물영역 검출, ii) 연

결 성분을 사용한 지침영역 숫자 분할, MSER을 사용한 기물영역
숫자 검출, iii) 숫자 영역 인식을 위한 CNN 기반의 인식. 제안하는
방법인 딥러닝 기반의 모델을 사용함으로써 다양한 계량기 타입에
강건하게 검출 및 인식이 가능하다. 따라서 검침원들이 기존처럼
숫자를 보고 일일이 입력할 필요없이 보다 능률적으로 일할 수 있
는 환경을 만들어 준다.

키워드: 가스 계량기 검침 시스템, 컴퓨터 비전, 이미지 프로세싱,
컨볼루션 뉴럴 네트워크(CNN)

Abstract

Meter Digit Detection and Recognition Method using YOLO v3 and CNN

Changeui Son

Major in Imaging Engineering-Digital Imaging

Department of Imaging Engineering

The Graduate School of Advanced Imaging Science, Multimedia & Film

Chung-Ang University

The meter reading-system field has been researched from conventional methods centered on image processing technology to techniques based on learning methods such as machine learning or deep learning. The biggest problem for meter reading systems based on computer vision is difficulty in recognizing the various kinds of meters. In fact, there are more than five major manufacturers for the meters installed in Korea. There are different meter reading areas, ID regions, and number formats by version. Because of these problems, most of the meter reading is still done hands-on. In this paper, we present an automatic meter-reading system that can work simply and efficiently, compared to existing meter reading systems that need a skilled worker. Our meter reading system consists of three parts: i) detection of

meter-reading and ID regions using You Only Look Once (YOLO), ii) digit segmentation for recognition, and iii) convolutional neural network (CNN)-based digit recognition. It is possible to robustly detect and recognize various meter types by using the method presented here. Therefore, it can provide an environment where gas meter checkers can work efficiently without inconvenient procedures.

Keywords: Gas Meter-Reading System, Computer Vision, Image Processing, Convolutional Neural Network.