동영상 추출 이미지파일을 이용한 머신러닝 검증 방법

곽경민^o, 노영주^{*}

^o한국산업기술대학교 스마트팩토라융합,

^{*}한국산업기술대학교 컴퓨터공학과
e-mail: k010511@naver.com^o, yrho@kpu.ac.kr^{*}

Verification Method for Machine Learning Based On Video Extraction ImageFiles

Ja-Sam Jeo^o, Ja-E Jeo^{*}

ODept. of Computer Engineering, Korea-Digital University,

*Dept. of Computer Science, Korea-Computer University

요 약 ●

이전 연구에 존재했던 영상에서 이미지를 추출하여 학습 데이터로 사용 할 때 시계열성을 고려하지 않은 상태에서의 검증은 정확하지 않을 수 있음을 설명한다. 정확한 형체를 가진 물체의 경우 매 프레임 마다 일 정한 모양을 유지할 가능성이 크지만, 기체나 액체처럼 유동성이 큰 형태를 분사 혹은 방류 할 때 순간적인 간섭 혹은 분산에 의해 실제 값이 분사 량 혹은 방류랑과 다를 수 있다. 본 연구에서는 이전 연구 중 Yolov3와 youtube 영상을 이용하여 연기 형태를 추출하고, 이를 Resnet에 학습시킨 연구를 이용하여 이와 비슷한 사례의 연구에서 나타날 수 있는 검증 오류들을 설명한다.

키워드: 머신러닝, 시뮬레이션

I Introduction

답라닝 기법은 2016년 3월 9일 이세돌과 알파고의 대결에서 답라닝 기반의 알파고가 승리를 하며 많은 관심을 받고 있다. 이후 많은 기업 및 연구자들이 딥러닝의 가능성을 발견하고, 이를 여러 분야에 접목사키기 시작한다. 머신러닝 기법은 이전의 간단한 퍼셉트론 모델 부터 Image recognition(이미지 인식)에 특화된 DenseNet까지 활발하게 연구 중이다. 이런 Image recognition에 있어 이전 연구들은 동영상 파일에서 이미지를 추출해 이를 인식 사키는 기술 들을 사용한다. 이때 사람 동물 특정 형태가 있는 물체들은 높은 인식률을 보이기도하지만, 기체와 액체와 같이 특정 형태가 없는 형태의 데이터는 정확히 인식하기가 힘들다. 또한 인식을 성공한다 하여도 순간적인 분산에의해 실제 값이 분사 랑과 같지 않을 수 있다.본 연구에서는 기체와액체와 같이 형태를 가지지 않는 물체에 대해 Image recognition 진행하여 검증 할 때 발생하는 오류들에 대해 설명하고, 검증 시고려할 상황에 대해 제시한다.

II Preliminaries

1. Related works

1.1 Yolo v3

Yolo v3는 Image recognition에 특화된 Darknet 기반의 어플라케이션이다. 이미지에 Boundary Box를 설정하는 과정을 거쳐 특정물체의 ROI(Region of Interest)를 설정하고, 이를 학습데이터로 Darknet에 학습 시켜 인식시키는 기술이다.[1] 기존의 Yolo와 다르게전체 데이터 정규화에 있어 정규분표 형태를 나타나지만, batch normalization 레이어를 추가하여, 학습 속도를 향상시키고, 인식률을 상승시켰다.

1.2 ResNet

ResNet은 ILSVRC 2015 image classification에서 우승한 신경망으로 image recognition에 특화 되어 있는 신경망이다.[2] 152개의 레이어로 구성되어 있는 심층 신경망으로 이전 연구에서는 마지막 레이어를 수정하여 사용하였다.

한국컴퓨터정보학회 하계학술대회 논문집 제28권 제2호 (2020. 7)

2. Previous research



Fig. 1. Experimental Data Collection Environment

<Fig 1> 과 같이 실험 환경을 구성하고, 연무기와 아크릴 파이프를 연결한다. 연무기 반대편 아크릴 파이프 끝부분에 마개를 만들어 마개에 구멍을 1개, 4개, 9개, 마개 없음 총 4가지의 실험을 진행한다. (이때 연무기의 무화 량은 일정하다.)

파이프 안에 갇혀 있는 연기는 수중기이기 때문에 압력에 의해 밖으로 추가적으로 분사되지 않고 파이프 내부에서 액화되므로 4가지 의 실험은 각각 육안으로 확인 가능한 차이를 보인다.

youtube 영상에서 추출한 연기 이미지를 Yolov3에 학습시키고, 학습 된 Yolov3 weight를 통해 실험 데이터의 연기 추출하고, ROI가 설정 된 연기 데이터는 resize 되어 다시 ResNet에 학습하여 무화량을 예측한다. <Fig. 2,3>은 Yolov3를 통해 ROI가 설정 된 예시와 100x100으로 resize 된 학습 데이터 예시이다.[3]



Fig. 2. Example of setting ROI



Fig. 3. Example of resized ROI

이렇게 힉습된 ResNet을 통해 얻어낸 구간별 오류율과 실데 데이터 와 예측 값은 아래의 <Fig 4,5>에서 확인 가능하다.

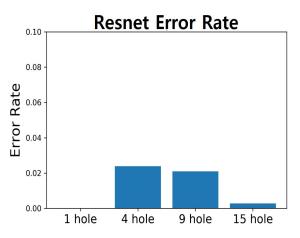


Fig. 4. Resnet error rate per section

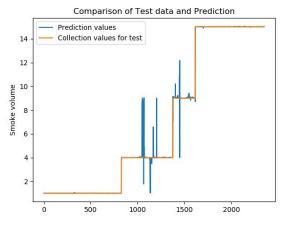


Fig. 5. Comparison with Test data and Prediction

III. The Proposed Scheme

2장에서 이전 연구의 방법과 결과에 대해 간략하게 살펴보았다. <Fig. 4>에 의하면 구간별 평균 error rate는 0.02를 넘지 않는다. 2%는 비교적 낮은 수치로도 보일 수 있으나 동영상과 같은 동적인

한국컴퓨터정보학회 하계학술대회 논문집 제28권 제2호 (2020. 7)

시계열성을 갖는 데이터는 정적인 CNN의 검증방법으로는 오류가 생긴다. 뒤에 첨부 될 figure들을 참고하면, 구멍의 개수가 많아질수록 연기가 짙어지고, 그로인해 하얀색 영역이 넓어져 오류의 원인은 히스토그램 분석으로 찾을 수 있다.

아래의 <Fig 6>은 <Fig. 5>에서 Prediction 값과 Collection data 값의 차이가 1.5이상 인 index만을 추출하여 해당 시진의 히스토 그램을 나타낸 것이다.

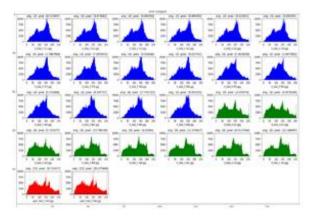


Fig. 6. Histogram of error section

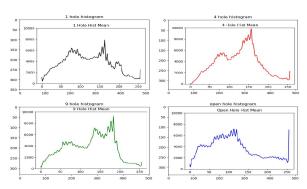


Fig. 7. Histogram means per each section

빨간색으로 그려진 그래프의 실제 값은 마개의 구멍이 4개 있는 경우이고, 초록색은 9개, 파란색은 마개 없음에 해당하는 영역이다. 아래의 <Fig. 7>은 각 마개의 구멍 개수 별로 Test set 히스토그램 평균이다.

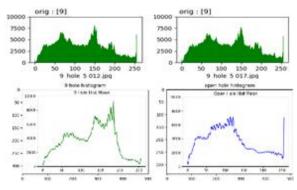


Fig. 8, ROI of Histogram

실제 우리가 관심을 가져야 할 부분은 <Fig. 8>과 같다. <Fig. 8>의 상단의 그래프 두개는 실제 분사 량은 9개의 구멍일 때이지만, 히스토그램의 모양은 하단의 두 그래프를 합친 것과 같이 나타난다.(하단의 두 그래프는 왼쪽부터 각각 마개 구멍 9, 마개 없음 이다.) 위와 같은 현상을 알아보기 위해서는 raw 데이터를 확인 해볼 필요가 있다.



Fig. 9. Comparison with Error and common raw data

<Fig. 9>의 상단 2개의 사진은 <Fig. 8> 상단 2개의 그래프의 raw 데이터이다. 하단의 2개의 사진은 일반적인 상황의 마개 구멍 9 경우의 사진이다. 상단과 하단의 사진 모두 같은 세기로 분사하였으나, 연기의 짙음은 육안으로도 확인이 가능하다.

위와 같은 현상으로 신경망에 학습을 시킬 때 실제 값은 true이나 false로 학습시키는 등의 오류가 발생 할 수 있다.

IV. Conclusions

본 연구에서 영상을 이미지로 추출하여 정적인 모델을 사용했을 때 발생할 수 있는 오류에 대해 다루었다. 이와 같은 오류는 신경망 자체의 성능을 자체를 저하시키거나, 학습 후 검증 할 때에 있어 잘못된 검증을 할 가능성이 있다.

최근 많은 연구가 이루어지고 있는 Ensemble 기법을 통해 정적 모델(CNN)과 동적 모델(LSTM)을 stacking 등의 기법을 사용해 각 모델의 단점을 보안이 가능할 것으로 판단된다. 이후 연구에서는 이러한 단점을 보안하고, 시계열성을 이용해 연기 확산의 동선을 파악하는 연구를 진행 할 예정이다.

REFERENCES

- [1] Kiana Ehsani, Hessam Bagherinezhad, Joseph Redmon, Roozbeh Mottaghi, Ali Farhadi, "Who Let The Dogs Out? Modeling Dog Behavior From Visual Data", CVPR 2018
- [2] Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun, "Deep Residual Learning for Image Recognition", LSVRC 2015
- [3] K.M Gwak, Young J. Rho, "Machine learning technique and image-based smoke emission prediction using Youtube video", KCC 2020 Busan