

# Daddy Makers

SW, HW, CG, ART, 건설, 건축 메이크 과정을 정리, 공유하는 블로그입니다 - 대디 메이커

2017년 6월 17일 토요일

## 딥러닝 Hello World - MNIST, CIFAR-10 데이터베이스 구조와 이미지넷

이 글은 딥러닝계의 Hello World 인 MNIST, CIFAR 데이터베이스 구조와 딥러닝 기술을 크게 발전시킨 세계 최대 이미지 학습 데이터 사이트 [이미지넷](#)을 소개한다.

### 1. MNIST 데이터베이스 구조

이 글은 [yann.lecun.com/exdb/mnist](http://yann.lecun.com/exdb/mnist) (Yann LeCun. NYU, Director of AI Research at Facebook, Silver Professor of Computer Science at the Courant Institute of Mathematical Sciences) 를 참고하였다. 이외에 많이 사용되고 있는 테스트 데이터셋은 다음과 같다.

- MNIST
- PTB(Penn Tree Bank)
- CIFAR-10, CIFAR-100

MNIST는 NYU의 Yann LeCun 사이트에서 다운로드 할 수 있다.

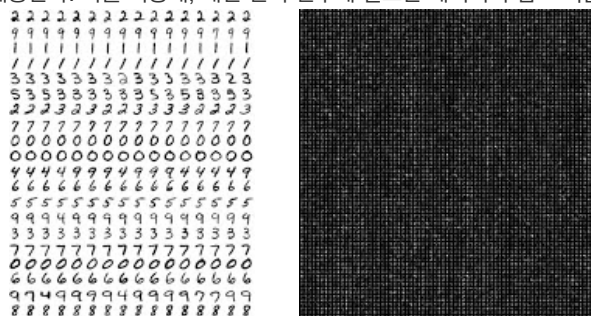
[train-images-idx3-ubyte.gz](http://train-images-idx3-ubyte.gz): training set images (9912422 bytes)

[train-labels-idx1-ubyte.gz](http://train-labels-idx1-ubyte.gz): training set labels (28881 bytes)

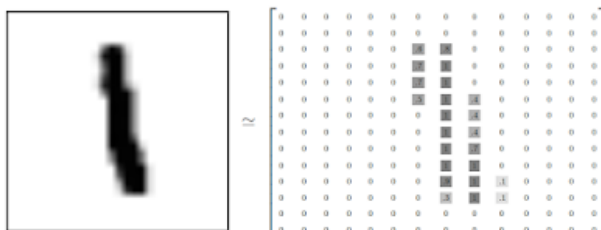
[t10k-images-idx3-ubyte.gz](http://t10k-images-idx3-ubyte.gz): test set images (1648877 bytes)

[t10k-labels-idx1-ubyte.gz](http://t10k-labels-idx1-ubyte.gz): test set labels (4542 bytes)

이 사이트에서는 딥러닝 신경망 학습에 필요한 필기체 숫자 이미지 60,000 훈련 집합, 10,000개 테스트 집합을 제공한다. 이를 이용해, 패턴 인식 연구에 필요한 데이터 수집 노력을 줄일 수 있다.



이 이미지는 원래 흑백이었지만, 28 x 28 픽셀 크기에 맞춰 정규화된 그레이 이미지로 처리되었다. 이를 위해, 안티 앨리어싱 기술을 적용하여, 이미지 픽셀 중심을 계산하여, 28x28 크기로 변환하였다.



MNIST 데이터는 SD-3와 SD-1이 있으며, SD3가 노이즈가 별로 없고 쉽게 인식할 수 있다. 참고로, SD-1은 500명의 사람이 쓴 58,527 이미지가 포함되어 있다.

프로필



Taewook Kang

G+ 팔로우 87

전체 프로필 보기

추천 게시물

### 키즈메이커 가이드 북 내용 소개



### DADDY MAKERS

우분투 운영체제 폴더 구조 및 핵심 명령어 - 2018. 3. 3.

우분투 백업 및 스냅샷 만들기 방법 소개 - 2018. 3. 3.

우분투 블랙 스크린 문제 해결 방법 - 2018. 3. 2.

이더리움 소스 빌드 및 스마트 계약 개발 소개 - 2018. 2. 27.

비트코인 소스 코드 빌드, 사용 및 블록체인 코드 구조 분석 - 2018. 2. 22.

전체 페이지뷰



149,645

구독



글



댓글

블로그 보관함

▶ 2018 (14)

▼ 2017 (32)

▶ 12월 (5)

▶ 9월 (3)

▶ 7월 (4)

▼ 6월 (8)

텐서플로우 기반 CIFAR 훈련모델 활용 이미지 인식

텐서플로우 기반 딥러닝 훈련 모델 파일 저장, 로딩 및 재활용

유니티 기반 3차원 포인트 클라우드 뷰어

딥러닝 Hello World - MNIST, CIFAR-10 데이터베이스 구조와 이미지넷

홀로렌즈 개발 환경 설정 및 앱 개발 방법

도커 기반 우분투 16.04 데스크탑 버전 설치

텐서플로우 기반 머신러닝 개념 및 사용 방법

오픈소스 기반 실시간 실내 위치 측정 시스템

▶ 5월 (4)

▶ 4월 (1)

MNIST 파일 형식은 MSB(Most Significant Bit. 가장 큰 숫자를 왼쪽에 기록하는 자리 표기법) 방식으로 저장된다. 각 파일의 형식은 다음과 같다.

TRAINING SET LABEL FILE (train-labels-idx1-ubyte):

[offset]	[type]	[value]	[description]
0000	32 bit integer	0x00000801(2049)	magic number (MSB first)
0004	32 bit integer	60000	number of items
0008	unsigned byte	??	label
0009	unsigned byte	??	label
.....			
xxxx	unsigned byte	??	label

라벨값은 0에서 9까지이다.

TRAINING SET IMAGE FILE (train-images-idx3-ubyte):

[offset]	[type]	[value]	[description]
0000	32 bit integer	0x00000803(2051)	magic number
0004	32 bit integer	60000	number of images
0008	32 bit integer	28	number of rows
0012	32 bit integer	28	number of columns
0016	unsigned byte	??	pixel
0017	unsigned byte	??	pixel
.....			
xxxx	unsigned byte	??	pixel

픽셀들은 row-wise로 구성되어 있다. 픽셀 값은 0에서 255이다. 0은 배경(흰색), 255는 전경색(검정색)을 의미한다.

TEST SET LABEL FILE (t10k-labels-idx1-ubyte):

[offset]	[type]	[value]	[description]
0000	32 bit integer	0x00000801(2049)	magic number (MSB first)
0004	32 bit integer	10000	number of items
0008	unsigned byte	??	label
0009	unsigned byte	??	label
.....			
xxxx	unsigned byte	??	label

라벨값은 0에서 9까지이다.

TEST SET IMAGE FILE (t10k-images-idx3-ubyte):

[offset]	[type]	[value]	[description]
0000	32 bit integer	0x00000803(2051)	magic number
0004	32 bit integer	10000	number of images
0008	32 bit integer	28	number of rows
0012	32 bit integer	28	number of columns
0016	unsigned byte	??	pixel
0017	unsigned byte	??	pixel
.....			
xxxx	unsigned byte	??	pixel

픽셀들은 row-wise로 구성되어 있다. 픽셀 값은 0에서 255이다. 0은 배경(흰색), 255는 전경색(검정색)을 의미한다.

IDX 파일 포맷은 간단한 벡터 형식으로 되어 있다.

magic number  
size in dimension 0  
size in dimension 1  
size in dimension 2  
.....  
size in dimension N  
data

▶ 2월 (2)

▶ 1월 (5)

▶ 2016 (47)

▶ 2015 (62)

태그

교육

IoT

메이커

머신러닝

미디어 아트

사물인터넷

로봇

비전

가상현실

드론

리눅스

항공

공간정보

그래픽스

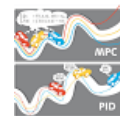
스마트시티

FOLLOW BY  
EMAIL

Email

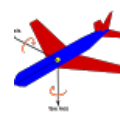
Submit

가장 많이 본 글



자율제어를 위한 PID 제어 개념 및 개발 방법

이 글에서는 PID 제어에 대해서 정리한다. PID는 액추에이터, 자율주행 차량, 로봇, 센서값 보정 등에 필수적인 함수로 사용된다. 1. 개요 PID제어는 목표한 물리량을 자연스럽게 수렴시키는 제어 방법이다. 우리는 PID제어를 본능적으...



IMU 센서 사용법

IMU센서 테스트 관련된 내용을 정리해 놓습니다. 참고로 아두이노, 프로세싱을 주로 사용하는 데, 이와 관련된 내용은 관련 사이트나 해당 도구 활용 사례를 정리한 레퍼런스를 참고하시길 바랍니다. 1. 개요 IMU는 inertial me...



텐서플로우 기반 딥러닝 훈련 모델 파일 저장, 로딩 및 재사용

텐서플로우를 사용해 딥러닝 훈련 모델 파일 저장 및 복구를 통한 재사용 방법을 간단히 기술한다. 딥러닝은 학습할 때 많은 시간이 소요된다. 학습된 모델을 저장해, 필요할 때 로딩하여 재사용하면, 이런 학습 시간을 생략할 수 있다. 1. softma...



웹기반 수학 교육 오픈소스 도구 GeoGebra

지오지브라(GeoGebra)는 2002년 오스트리아 마르쿠스 호헨바터에 의해 개발된 오픈소스 기반 수학 교육 도구이다. 지오지브라는 기하, 대수, 미적분, 통계 및 이산수학을 쉽게 다룰 수 있다. 아이패드, 안드로이드, 윈도우 8 앱에서 실행되며,

매직 넘버는 정수형(MSB)이다. 첫번째 2바이트는 항상 0이고, 세번째 바이트 코드는 다음을 의미한다.

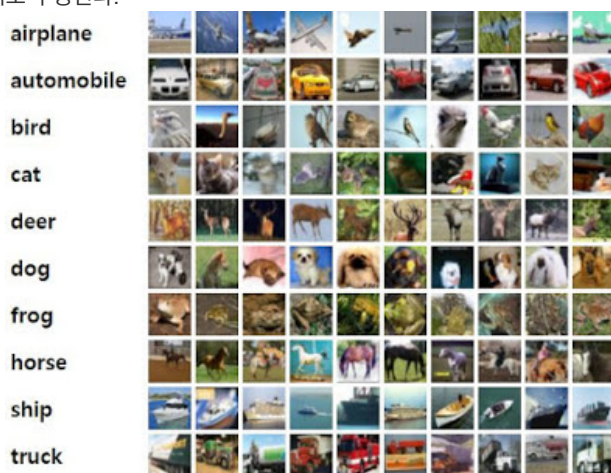
0x08: unsigned byte  
0x09: signed byte  
0x0B: short (2 bytes)  
0x0C: int (4 bytes)  
0x0D: float (4 bytes)  
0x0E: double (8 bytes)

## 2. CIFAR 데이터베이스 구조

CIFAR-10과 CIFAR-100은 80 백만개의 소형 이미지 데이터셋이다. 이 데이터는 Alex Krizhevsky, Vinod Nair, Geoffrey Hinton이 모았다. 이 글은 [Krizhevsky 홈페이지](#)를 참고하였다. 이 이미지를 이용해, 다양한 머신러닝 기법이 테스트되고 있다(예. [텐서플로우 기반 CNN 이미지 훈련](#))

CIFAR-10은 60000 32x32 컬러 이미지로 10개 클래스로 구성된다. 50000개의 훈련 이미지와 10000개의 테스트 이미지가 있다.

데이터셋은 5개 훈련 배치 셋과 한개 테스트 배치 셋으로 구분되며, 각 셋은 10000 개 이미지이다. 테스트 배치 셋은 1000 개 선택된 임의 이미지가 포함되어 있다. 훈련 배치 셋은 각 클래스별로 5000개 이미지로 구성된다.



다음은 다운로드 링크이다.

Version	Size	md5sum
<a href="#">CIFAR-10 python version</a>	163 MB	c58f30108f718f92721af3b95e74349a70270af85842c9e89bb428ec9976c926c32a1d4
<a href="#">CIFAR-10 Matlab version</a>	175 MB	
<a href="#">CIFAR-10 binary version (suitable for C programs)</a>	162 MB	

## 3. 이미지 넷

이미지넷은 Fei-Fei Li 교수가 2007년 제안한 학습용 이미지 데이터베이스 구축 아이디어에서 시작된 프로젝트이다. 비전 기술 개발 시 필요한 세계 최대 학습용 이미지 데이터베이스가 무료로 제공

HT...



**비트코인 소스 코드 빌드, 사용 및 블록체인 코드 구조 분석**

얼마전 BIM principle에 올린 블록체인과 BIM - 스마트 계약을 위한 블록체인 기술에 대한 기고 요청이 있었다. 이 글은 4차 산업혁명의 핵심 키워드로 알려진 블록체인 응용인 비트코인 정체와 내부 메커니즘을 좀 더 깊이 확인하고 설명하기

...



**텐서플로우 기반 머신러닝 개념 및 사용 방법**

이 글은 텐서플로우 학습 방법을 요약한다. 이 글은 텐서플로우(버전 1.1)를 이미 설치하였다고 가정하고 진행한다. 여기서 실행된 샘플은 아래 e-book 레퍼런스에서 제공된 것이다. 이 레퍼런스들은 구글 텐서플로우 설치때 함께 설치되는 예제들을 기반...



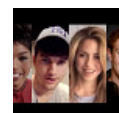
**텐서 플로우 및 케라스 최신버전 설치 방법 (Windows 10)과 개념**

이 글은 텐서플로우(Tensorflow) 1.1, 케라스(keras)를 윈도우 10 환경에서 설치하는 순서를 정리한다. 기타, 주피터 노트북에 사용되는 NumPy 같은 파이썬 라이브러리와 케라스 설치도 함께 다루겠다. 참고로, 현재(2018.2.8...



**아두이노 기반 6 DoF 로봇암 조립 방법 및 제어**

오늘은 6축의 자유도를 가진 6 DoF (자유도) Robot Arm Kit 조립 및 개발 방법을 정리해 본다. 6축 자유도는 사람의 팔 구조와 매우 유사한 동작을 할 수 있다. 사용할 로봇암은 Oak Studio 6 DoF 이다. 여기에 사용...



**쉽고 저렴한 어린이 용 코딩 교육 소개**

몇년 전부터 코딩 교육에 대한 이야기가 많아지고 있습니다. 코딩 교육 소개 영상 이 영상처럼, 아이들에게 코딩을 통한 논리적 사고 능력, 창의력을 키워주고 싶은데, 방법이 마땅히 떠오르지 않는 경우가 있습니다. 이럴때 학원을 알아보는 경...



**라즈베리파이 기반 네트워크 감시 카메라**

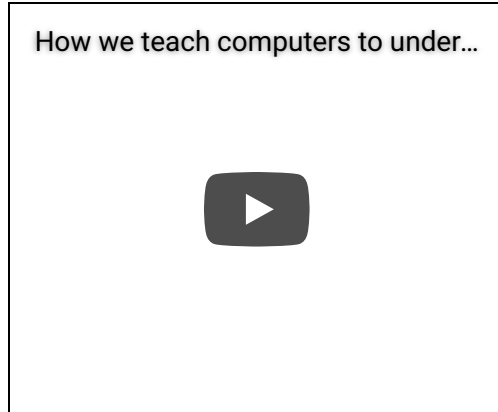
본 글은 라즈베리파이 기반 감시 카메라를 만드는 방법을 간략히 정리한다. 네트워크로 연결된 감시카메라는 최소 6~7만원, 모션 감지 등 고급 기능이 지원되는 제품은 몇 십만원 이상 넘어간다. 하지만, motion, motionPie(motion p...

된다.

이미지넷은 무료로 학습용 데이터베이스를 제공하고 있다. 이미지넷 초창기 Fei-Fei Li 와 주변 동료들은 약 15,000,000 이미지에 대한 라벨링 작업을 큰 R&D펀드 없이 진행하였으며, 이후 도움을 준 세계 각국의 연구자들도 그녀의 아이디어에 영감을 받아, 개인의 시간과 자원을 공헌하였다.

이미지넷은 매년 비전 기술의 사물 인식 정확도를 기준으로 평가하는 대회를 열고 있다. 대회에서 우수한 알고리즘을 평가된 기술은 큰 주목을 받고, 오픈소스로 공개되어, 많은 사람들에게 도움을 주고 있다. 2012년 캐나다 토론토 대학의 [알렉스 크리제스브키](#)는 GPU 기반 CNN 딥러닝 모델 ([ImageNet Classification with Deep Convolutional Neural Networks, 2012](#))을 이용해, 80% 이상 정확도를 보여주어, 사람들을 놀라게 했다. 그는 이 기술을 오픈소스로 공개하였으며, 마이크로소프트는 이 기술을 기반으로 정확도를 96%까지 끌어올렸다([관련 기사](#)).

이미지넷은 비전 분야에서 오픈소스를 기반으로 공유하고 발전하는 문화를 만들게 된 큰 계기가 되었다. 이와 관련된 자세한 내용은 다음 Fei-Fei Li 교수의 TED 강연을 통해 확인할 수 있다.



ImageNet (Fei-Fei Li 교수. TED)

최근 스탠포드대학에서 구글로 자리를 옮긴 Fei-Fei Li 교수에 대한 좀 더 자세한 내용은 다음 링크에서 살펴볼 수 있다.

- Fei-Fei Li 교수([CNN 인터뷰](#), [Best China News 인터뷰](#))

#### 4. 마무리

이 글에서 머신러닝 훈련에 필요한 데이터를 제공하는 레퍼런스를 간략히 살펴보았다. 이외에, 이 분야에서 공헌한 연구자 중 한명인 Fei-Fei Li 교수의 ImageNet에 대해서도 간단히 소개해 보았다. 이외에 데이터셋이 필요한 경우, [Kaggle](#) 등을 방문해 학습에 필요한 데이터를 얻을 수 있다.

작성자: [Taewook Kang](#) 시간: 오전 6:09



댓글 없음:

댓글 쓰기

댓글을 입력하세요...



작성자

Unknown (Goo-)

로그아웃

게시

미리보기

☐ 알림