# CNN을 사용한 이미지 기반의 안드로이드 멀웨어 패밀리 분류

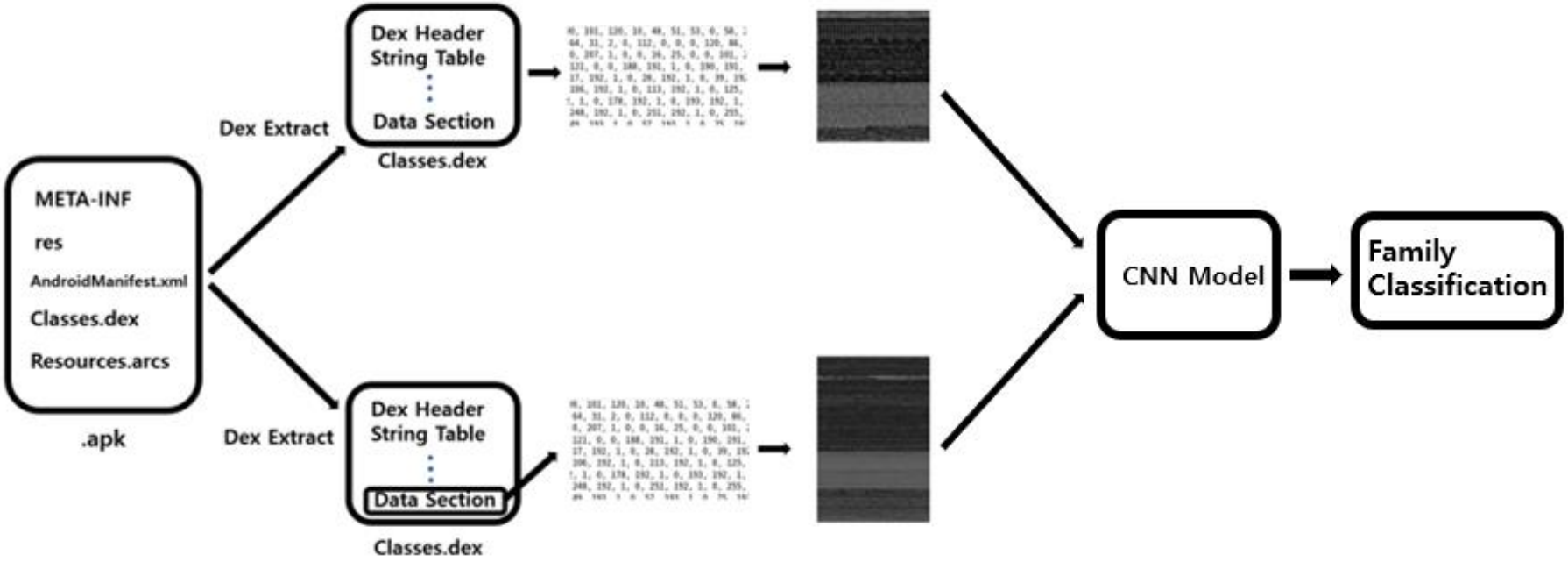
Image-based Android Malware Family Classification Using Convolutional Neural Network

2020 KCC

강문영

loveskangi@gmail.com
Dept. of Software Science
Dankook University

# 시스템 구성

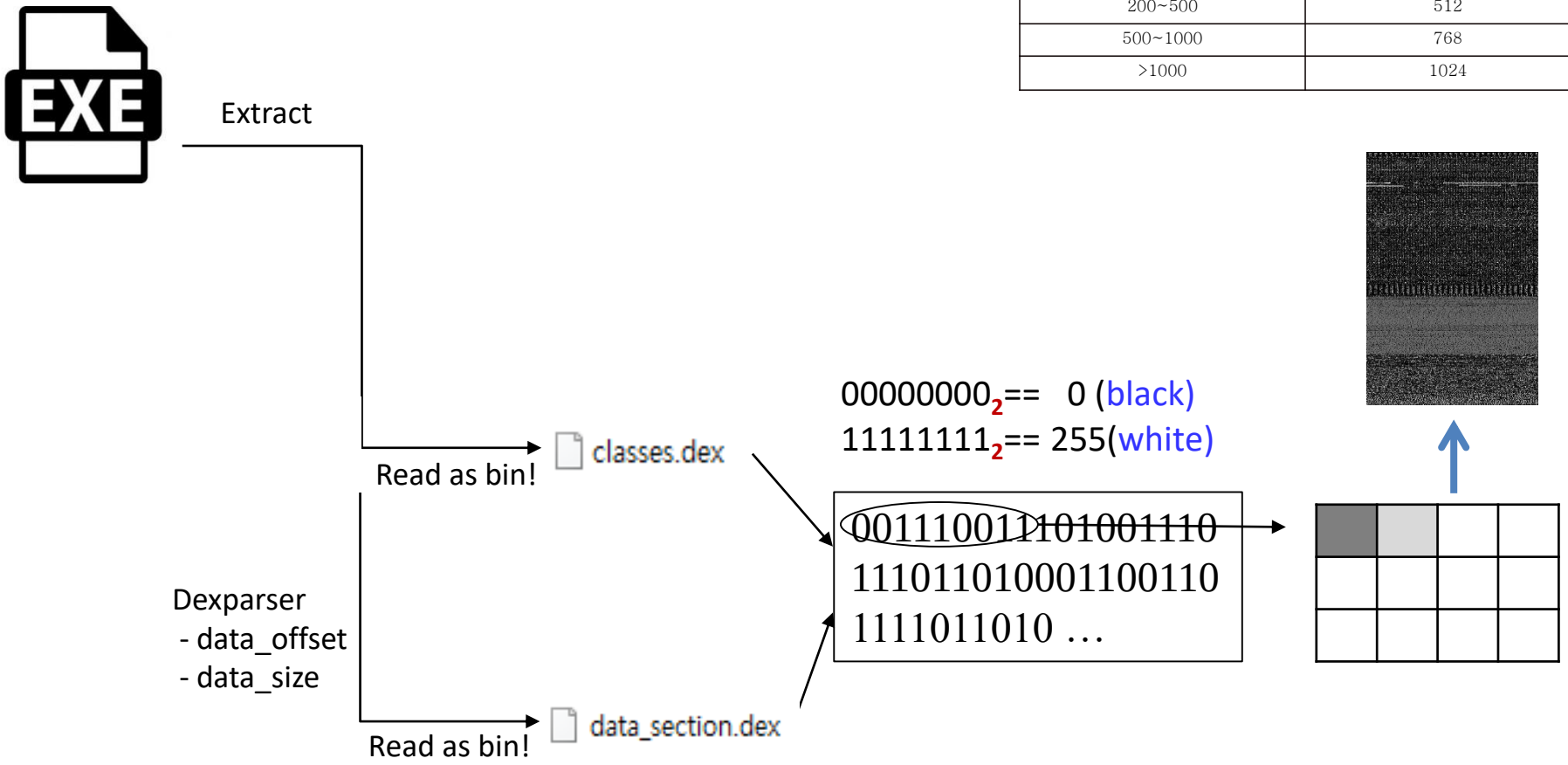



- Dex 파일 기반의 Gray Scale Image CNN 실험
- Data_Section 파일 기반의 Gray Scale Image CNN 실험

# 악성 앱 이미지 생성

- 악성 앱 이미지 생성
- APK 파일 압축 해제 Classes.dex파일 추출

| File Size Range(KB) | Image Width(pixel) |
|---|---|
| <10 | 32 |
| 10~30 | 64 |
| 30~60 | 128 |
| 60~100 | 256 |
| 100~200 | 384 |
| 200~500 | 512 |
| 500~1000 | 768 |
| >1000 | 1024 |



# Dataset

| Family | Num. of samples |
|---|---|
| BaseBridge | 326 |
| DroidDream | 81 |
| DroidKungFu | 666 |
| FakeDoc | 132 |
| FakeInstaller | 921 |
| FakeRun | 61 |
| Gappusin | 58 |
| Iconosys | 152 |
| Imlog | 43 |
| Kmin | 147 |
| MobileTX | 69 |
| Opfake | 601 |
| Plankton | 624 |
| Sendpay | 59 |

- 총 14개의 패밀리 실험
- Training set : 8 / Test set : 2

# 성능 평가

## 실험 1 Dex vs Data_Section

| | Dex | Data Section |
|---|---|---|
| Accuracy | 97% | 95% |
| Average File size | 164KB | 134KB |

Average 18% Gap!

### Dex 파일을 이용한 결과

| Family | Precision | Recall | F1-score |
|---|---|---|---|
| BaseBridge | 0.91 | 0.92 | 0.92 |
| DroidDream | 0.86 | 0.75 | 0.80 |
| DroidKungFu | 0.96 | 0.95 | 0.96 |
| FakeDoc | 0.90 | 1.00 | 0.95 |
| FakeInstaller | 0.99 | 0.99 | 0.99 |
| FakeRun | 0.80 | 1.00 | 0.89 |
| Gappusin | 1.00 | 0.75 | 0.86 |
| Iconosys | 1.00 | 0.97 | 0.98 |
| Imlog | 0.89 | 0.89 | 0.89 |
| Kmin | 1.00 | 1.00 | 1.00 |
| MobileTX | 1.00 | 1.00 | 1.00 |
| Opfake | 0.98 | 0.99 | 0.99 |
| Plankton | 0.98 | 0.98 | 0.98 |
| Sendpay | 1.00 | 1.00 | 1.00 |

### Data_Section 파일을 이용한 결과

| Family | Precision | Recall | F1-score |
|---|---|---|---|
| BaseBridge | 0.96 | 0.78 | 0.86 |
| DroidDream | 0.88 | 0.94 | 0.91 |
| DroidKungFu | 0.91 | 0.97 | 0.94 |
| FakeDoc | 0.92 | 0.92 | 0.92 |
| FakeInstaller | 0.96 | 0.99 | 0.97 |
| FakeRun | 0.92 | 1.00 | 0.96 |
| Gappusin | 0.78 | 0.58 | 0.67 |
| Iconosys | 0.88 | 0.94 | 0.91 |
| Imlog | 1.00 | 0.89 | 0.94 |
| Kmin | 0.96 | 0.93 | 0.95 |
| MobileTX | 1.00 | 1.00 | 1.00 |
| Opfake | 0.97 | 0.95 | 0.96 |
| Plankton | 0.98 | 0.98 | 0.98 |
| Sendpay | 1.00 | 1.00 | 1.00 |

# 성능 평가

## 실험 2 기존 연구와 비교

| Family | Precision | Recall | F1-score |
|---|---|---|---|
| BaseBridge | 0.96 | 0.84 | 0.90 |
| DroidDream | 0.88 | 0.83 | 0.85 |
| DroidKungFu | 0.73 | 0.93 | 0.82 |
| FakeDoc | 0.98 | 0.96 | 0.97 |
| FakeInstaller | 0.95 | 0.97 | 0.96 |
| FakeRun | 1.00 | 0.95 | 0.97 |
| Gappusin | 0.89 | 0.96 | 0.92 |
| Iconosys | 1.00 | 0.83 | 0.91 |
| Imlog | 0.97 | 0.98 | 0.97 |
| Kmin | 0.99 | 1.00 | 0.99 |
| MobileTX | 0.95 | 0.97 | 0.96 |
| Opfake | 0.97 | 0.97 | 0.97 |
| Plankton | 0.95 | 0.92 | 0.93 |
| Sendpay | 1.00 | 0.53 | 0.69 |

# 결론 및 향후 연구

- 결론
  - CNN을 이용한 악성 앱의 패밀리 분류
    - Dex 파일과 Dex 파일 내부의 Data Section만을 분리하여 이미지 생성
    - 18% 차이가 발생하는 2가지의 파일을 이용한 비교 실험 진행
    - 2% 의 정확도 차이 발생, Droid Dream 과 같은 특정 패밀리에서

    성능 지표간 차이가 간소화 및 수치 향상 하지만 Gapppusian은 성능지표가 낮음

- 향후 연구 방향
  - CNN과 LSTM을 이용한 다중입력 모델 실험
    - Permission 정보와 API 호출을 이용한 LSTM 실험
    - 기존 CNN과의 가중치 값을 공유하는 다중입력 모델 구현
    - 이외의 다양한 모델에서 실험 진행

# Thank you