Predicting Seismic-induced Building Damage using Mixture Deep Learning Features

Stanford University

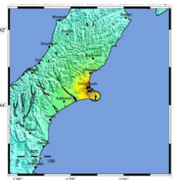
Zhaozhuo Xu Zhiyuan Li Haiwen Wang
{zhaozhuo, zhiyuan8, hwwang} @stanford.edu

CS 230 Deep Learning

Motivation

Earthquake building damage prediction models can provide necessary information for urban planer and insurance decision-makers. In this project:

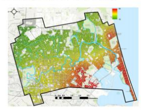
1) Trained a CNN to obtain geotechnical feature vectors.
2) Consturcted mixture features via combining geotechnical vectors with building information matrix,
3) Used ANN and softmax to predict damage degree.
4) Outperform currently statistical methods and National technical specification.

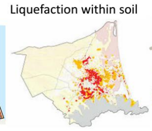


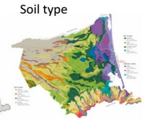
Dataset Description

Features

- Earthquake information[*1]
  - **Pgv** Peak ground velocity
  - **Liq** Liquefaction within soil
  - **Soil** Soil type
- Building information[*2]
  - **year:** building year
  - **floor:** floor area
  - **storey:** number of storeys
  - **incurred:** estimated loss
  - **replace:** dwelling value
  - **occu:** occupancy type (residential, commercial)
  - **type:** wall and roof material
  - **damage ratio (Dr):** , 4 categories





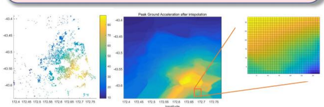


[*1] NZ GNS Science. https://lris.scinfo.org.nz/
[*2] Housing and property. NZ Government. https://www.govt.nz/

Approaches

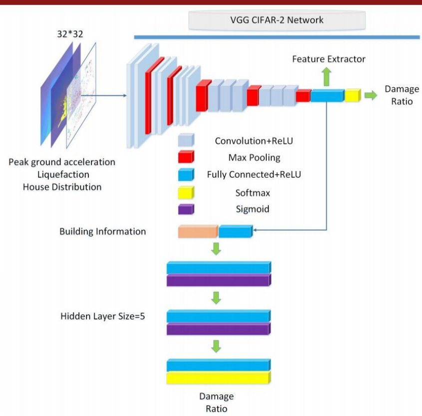
Constructing CNN inputs

- The earthquake information (pgv/liq/soil) is from observation recordings (see figure at left).
- Using KNN, datasets are interpolated as 5000×5000 huge picture (see figure in middle)
- Earthquake information around buildings is saved as 32×32 tiles. (see figure below)



Mixture DL Feature

- Refine a VGG CIFAR-2 network for two class classification
- Construct 2 class categorical damage state label
- Train VGG network
- Obtain the FC layer as features extracted from input tiles.



Model Parameter

CNN (vgg network)	Epoch	100	ANN	Input	10 + 2	Learning rate	0.01
	Batch-size	128		$n^{[1]}$	10	Batch-size	64
	Learning rate	0.02		$n^{[2]}$	5	Epoch	1000
	Output	softmax		Optimizer	Adam	Activation	sigmoid

Train and Test Set Construction

- Each data point contains: pgv, liq, soil, year, floor, story incurred, replace, occu, type
- Label for each data point: damage ratio
- Randomly shuffle 6788 data points
- Split into training set(6000 points) and test set(788 points)

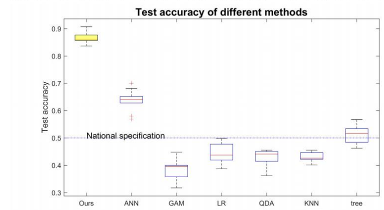
Labels and Loss Function

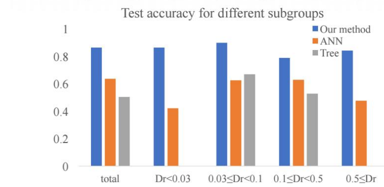
- Damage Ratio (Dr) = $\frac{\text{Repair (repair cost)}}{\text{Replace (dwelling value)}}$
- Damage degree:
  Small: Dr < 0.03, Middle: 0.03 < Dr < 0.1
  Large: 0.1 < Dr < 0. Extra Large: Dr > 0.5
- Loss function: softmax cross entropy
  $$\text{Minimize } \sum_{i=1}^{\#\,classes} - y_i \log g(\hat{y}_i)$$
- Feature extractor (VGG) training:
  # classes = Damage State
  labeled as: 0,1
- Analytical neural network training:
  # classes = Damage degree
  labeled as: 0,1,2,3
- Evaluation metric: categorical predicting accuracy

References:
[1] BA Bradley. Ground motions observed in the Darfield and Christchurch earthquakes and the importance of local site response effects[J]. New Zealand Journal of Geology & Geophysics, 2012, 55(3):279-286.
[2] Lautour O R D, Omenzetter P. Prediction of seismic-induced structural damage using artificial neural networks[J]. Engineering Structures, 2009, 31(2):600-606.

Results

(1) After 10 trials the combined CNN model could achieve 86.5% mean testing accuracy, while ANN model without information from CNN reaches 65.2% mean testing accuracy. Compared to other statistical methods and national specification method, our model outperforms greatly.
(2) Our model possesses 86.5%, 90.0%, 78.9%, 84.3% accuracy for damage ratio group 0-0.03, 0.3-0.1, 0.1-0.5, 0.5-∞, respectively. It has stably excellent performance for each subgroup in contrast with others.





Conclusion

By combining the training of CNN and ANN models, we manage to find sets of DL features that are able to yield an excellent damage ratio estimation. To be specific:

- We develop a ANN that outperform other state-of-art statistical methods. This is also the first CNN model applied to the specific area.
- Features extracted by CNN can help improve the performance of ANN when combining building information.
- The most important three features are 'replace (dwelling value)', 'floor (floor area)', 'occu (occupancy type)'. (figures are omitted)

For future works:

- Tuning a better CNN model.
- Use this feature in other damage ratio based tasks..