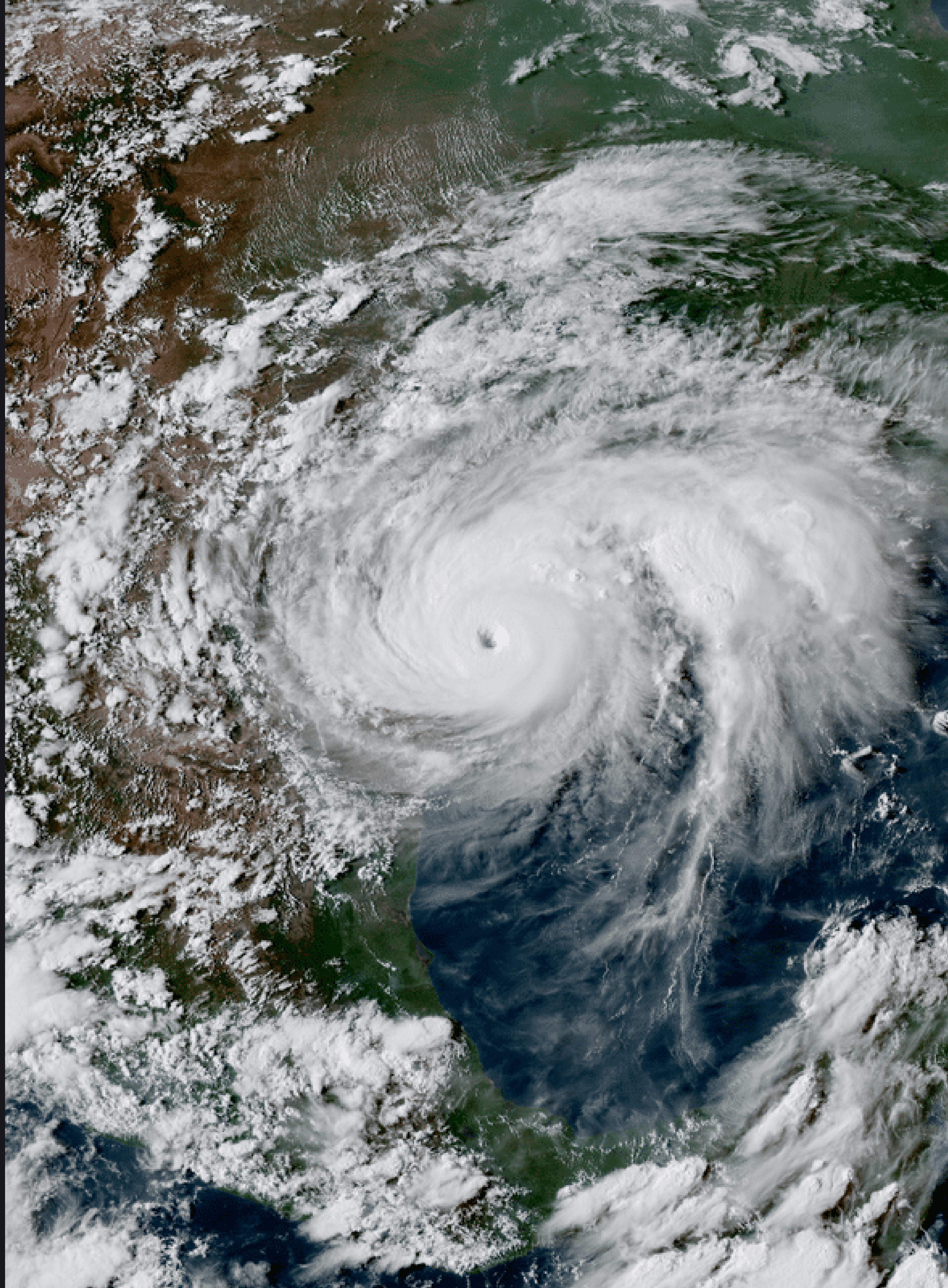
DETECTING DAMAGED BUILDINGS ON POST-HURRICANE SATELLITE IMAGERY

AN VO QUANG | JANUARY 2020 | HACKSHOW

2017 HURRICANE HARVEY

- Formed on August 17 and dissipated on September 2
- Category 4 hurricane that made landfall on Texas and Louisiana
- The wettest tropical cyclone on record in the United States.



30 000 DISPLACED PEOPLE

Harvey damaged 200,000 homes. More than 500,000 people asked for federal assistance.

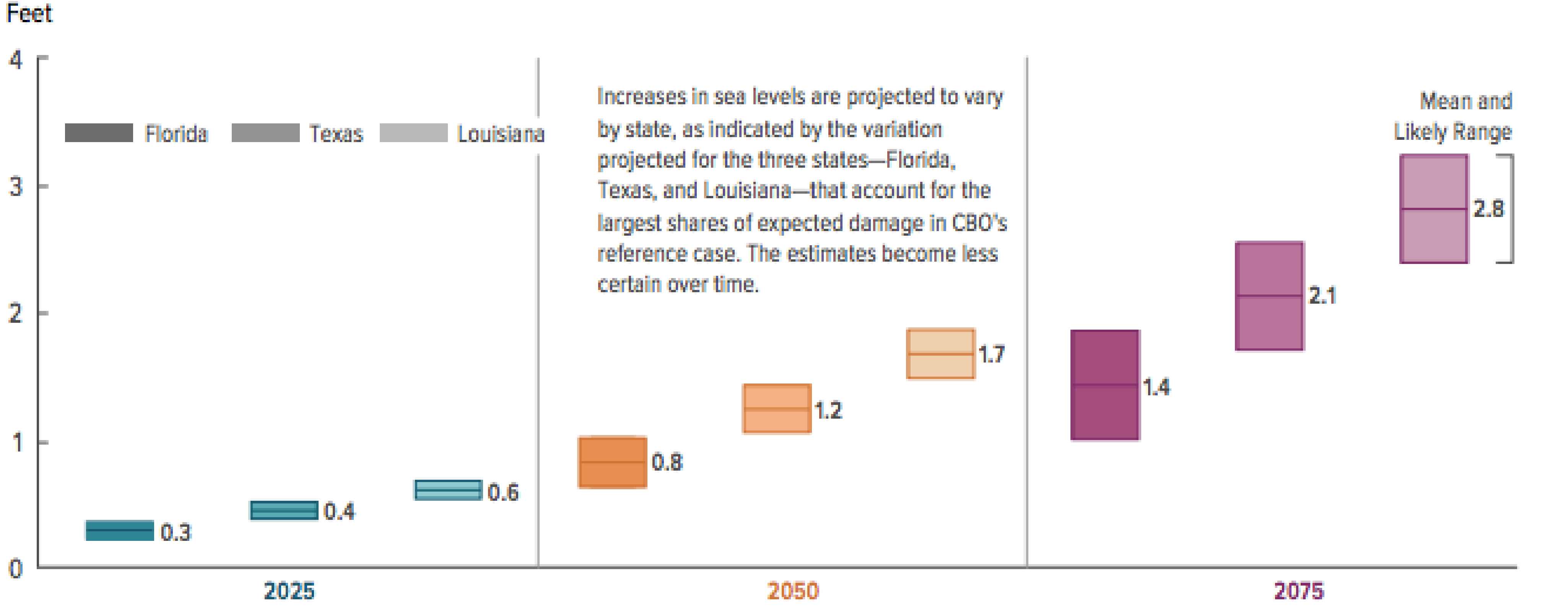
$125 BILLION

2nd costliest tropical cyclone, from catastrophic rainfall-triggered flooding in Texas, Louisiana, Missouri and Tennessee.

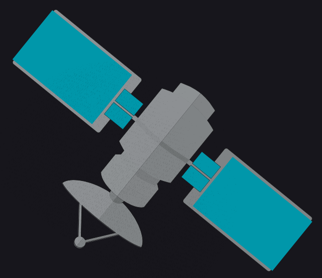
M.I.T. models predict more hurricanes by 2035 and 11% will be Category 3, 4, and 5

- 1.2 million Americans live in coastal areas at risk of substantial damage from hurricanes.
- Government costs for hurricane damage are $28 billion a year and will increase to $39 billion by 2075
- Causes: climate change (global warming means higher ocean temperatures at deeper depths to feed hurricane strength).

Projected Rise of Sea Levels in Florida, Texas, and Louisiana in Selected Future Years



Source: Congressional Budget Office, using data from Risk Management Solutions; Robert E. Kopp and others, "Probabilistic 21st and 22nd Century Sea-Level Projections at a Global Network of Tide-Gauge Sites," *Earth's Future*, vol. 2, no. 8 (August 2014; corrected, October 2014), http://onlinelibrary.wiley.com/doi/10.1002/2014EF000239/full.



WHY USING SATELLITE IMAGES?

Damage assessment is critical to emergency managers and first responders so that resources can be planned and allocated appropriately.

Satellite images have wider fields of view than ground search and aerial survey and can avoid injury risk during search and rescue.

OBJECTIVE

Make a model able to automatically identify if a given region is likely to contain flooding damage

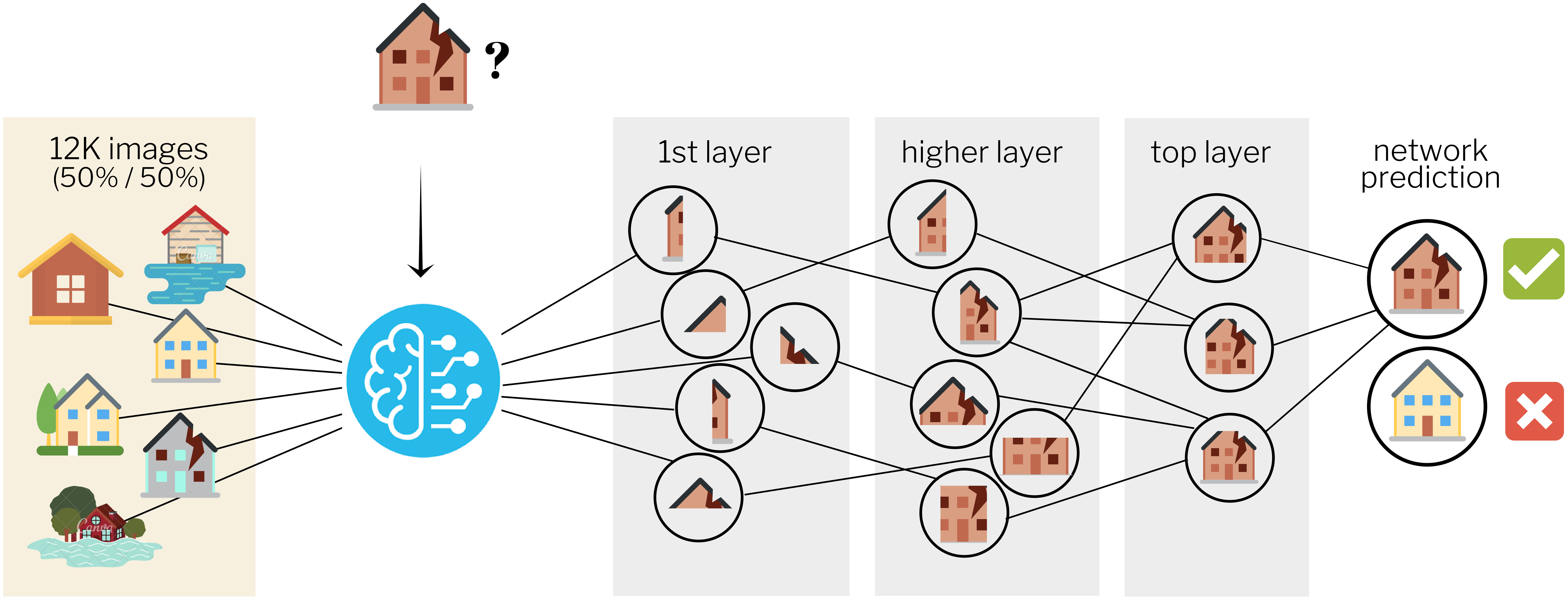
The dataset DigitalGlobe™

12 000 square-sized images extracted from DigitalGlobe satellite imagery from Texas after Hurricane Harvey. Data can be found on Kaggle.

The labels

Images were labeled by volunteers in Tomnod crowd-sourcing project as Damage/flooded or No damage.

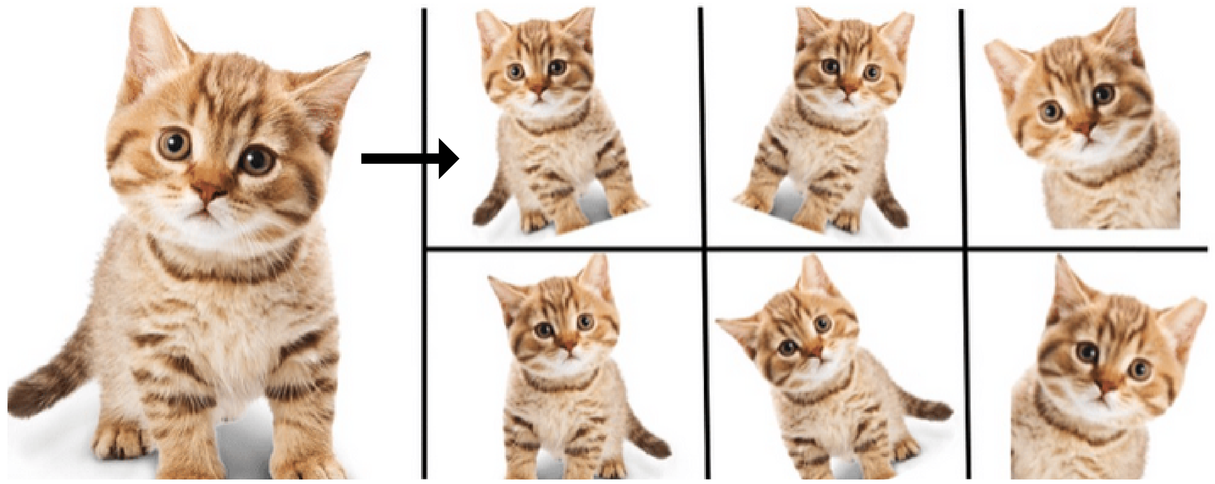
HOW NEURAL NETWORKS RECOGNIZE A DAMAGED HOUSE IN A PHOTO



MY APPROACH

IMAGE PREPROCESSING

Data augmentation, rescaling

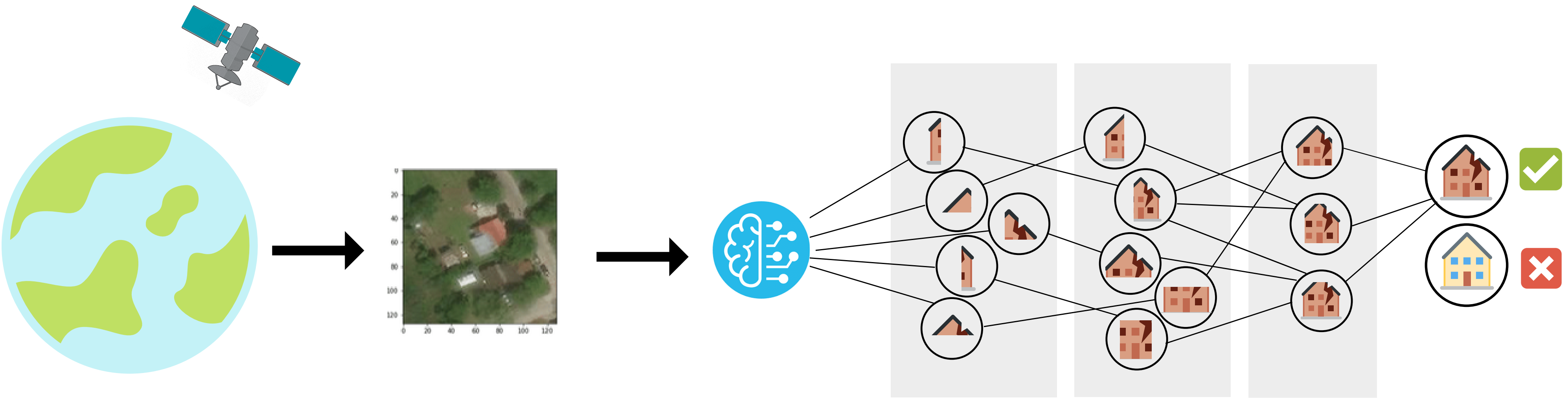


MODEL TESTING

Using Convolutional Neural Networks (CNN)

PREDICTIONS WITH SELECTED MODEL

2000 images in testing data

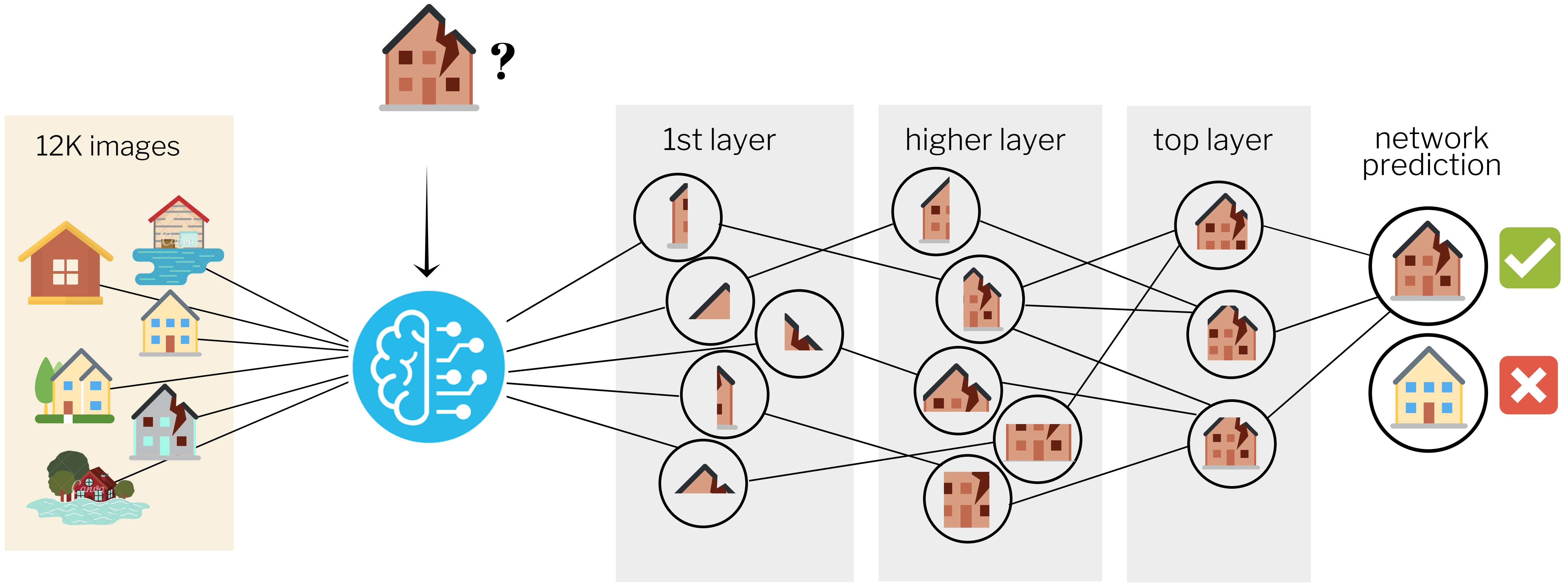


FIRST DESIGNED MODEL

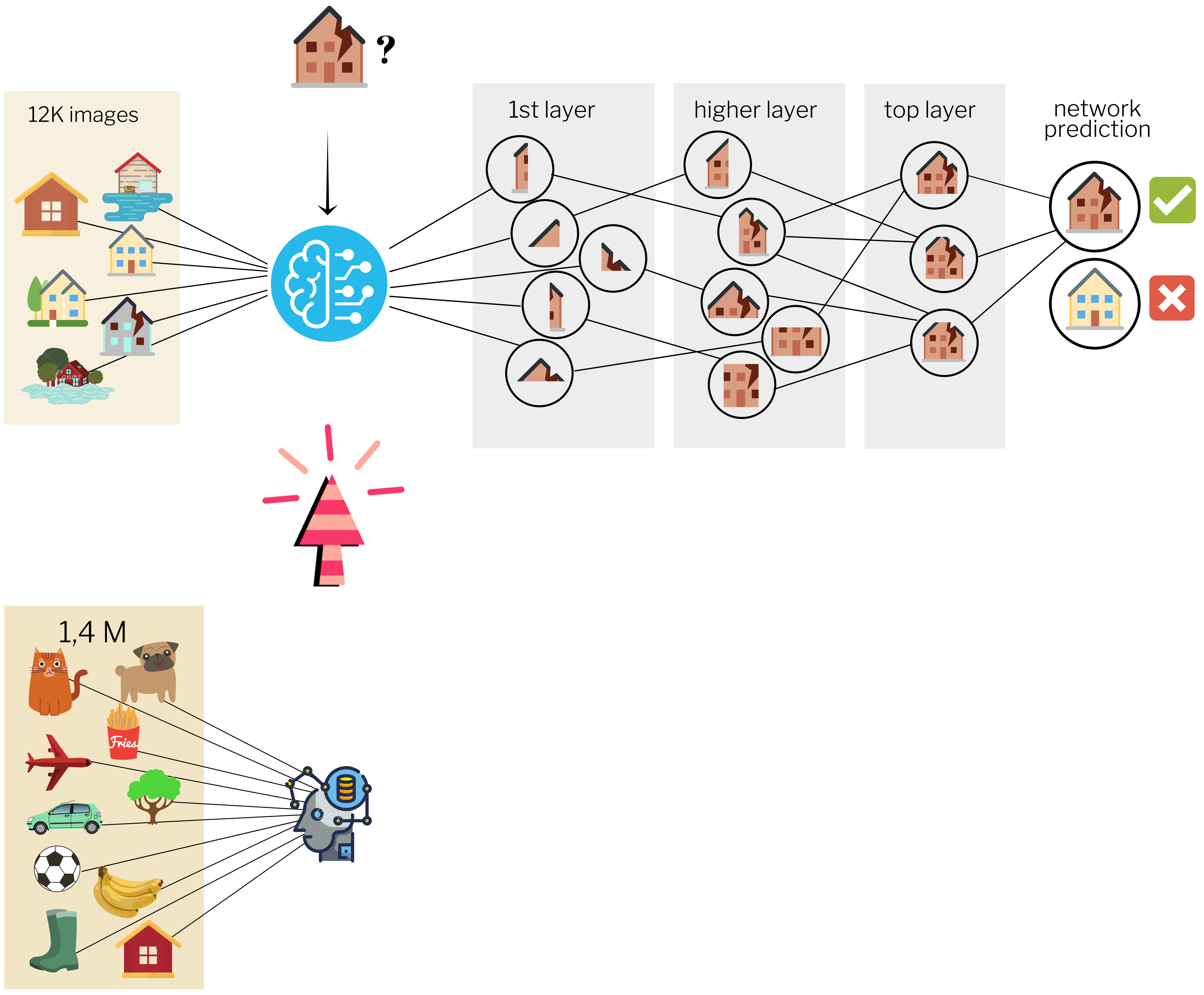
ACCURACY = 97.1%

1942 WELL CLASSIFIED HOUSES

58 WRONG CLASSIFICATION



TRANSFER LEARNING

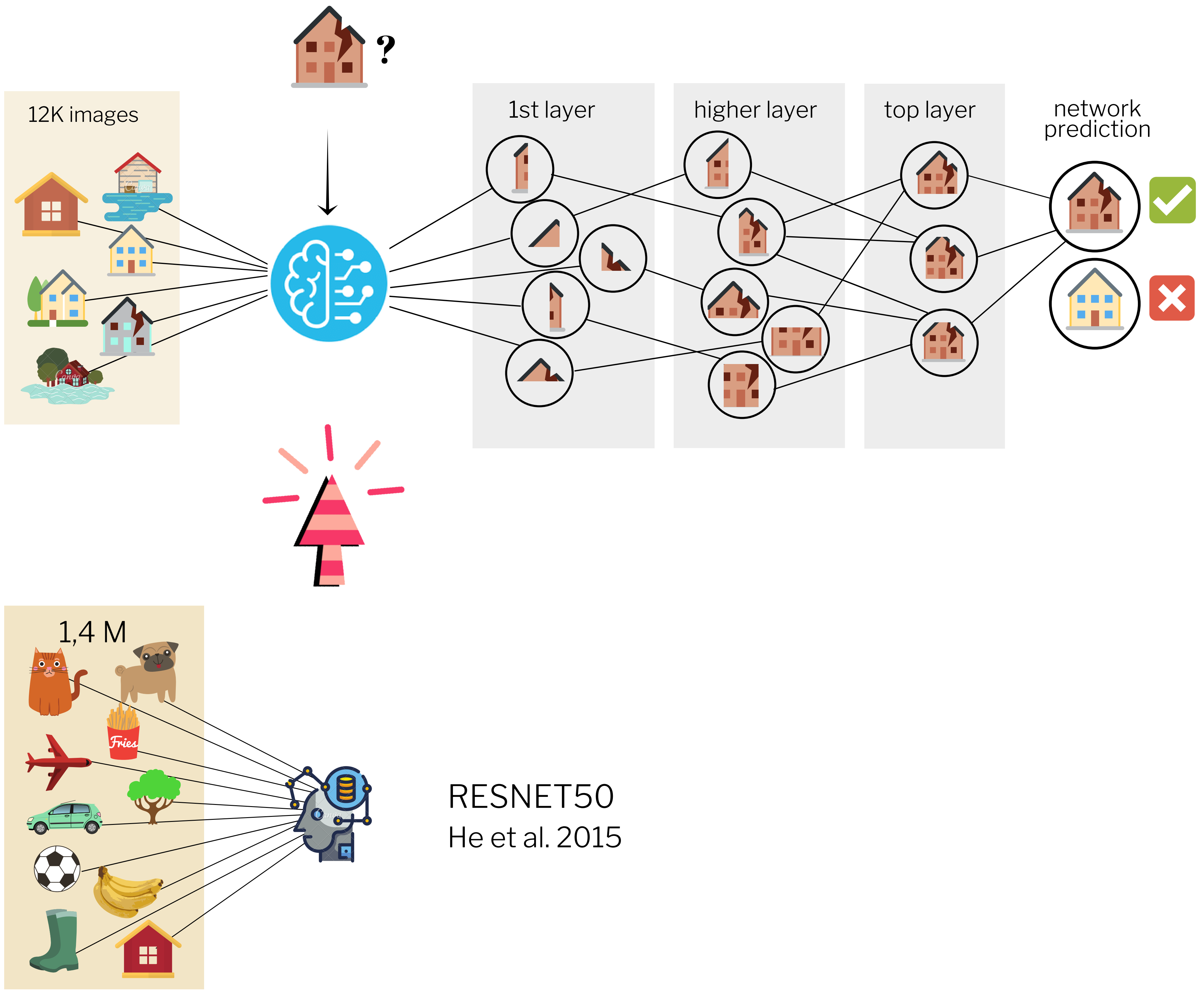


FINAL MODEL

ACCURACY = 99.6%

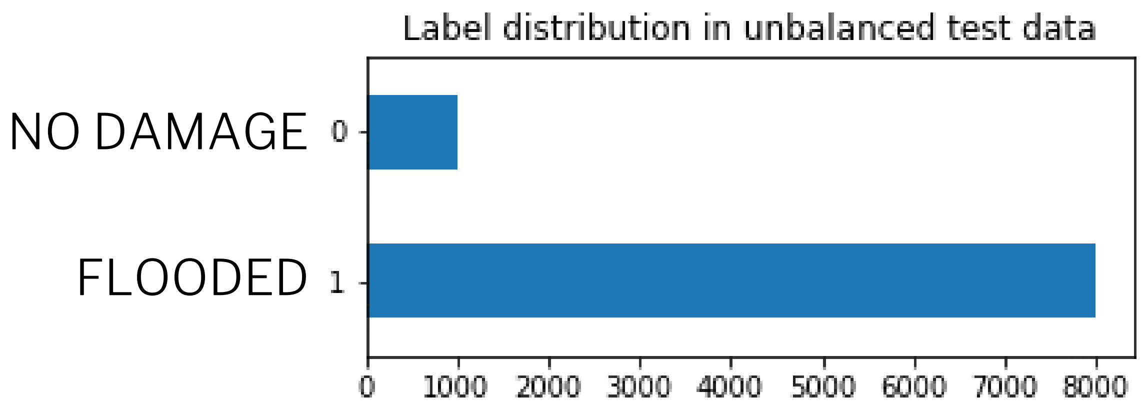
✓ 1992 WELL CLASSIFIED HOUSES

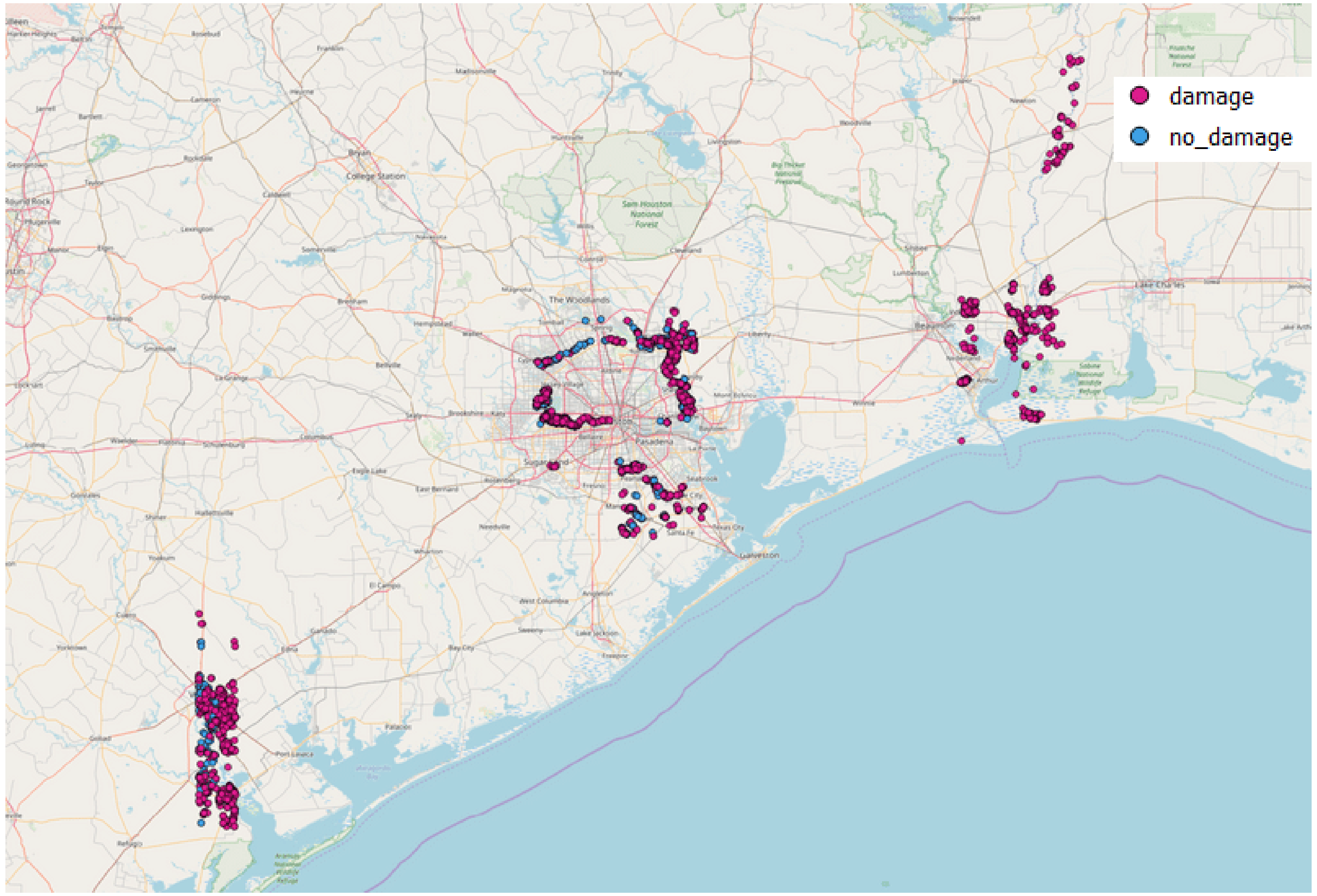
✗ 8 WRONG CLASSIFICATION



RESNET50
He et al. 2015

PERFORMING OUR MODEL ON LARGER UNBALANCED DATA





9000 HOUSES IN HOUSTON AREA

✅ 8943 GOOD CLASSIFICATION

❌ 57 WRONG CLASSIFICATION

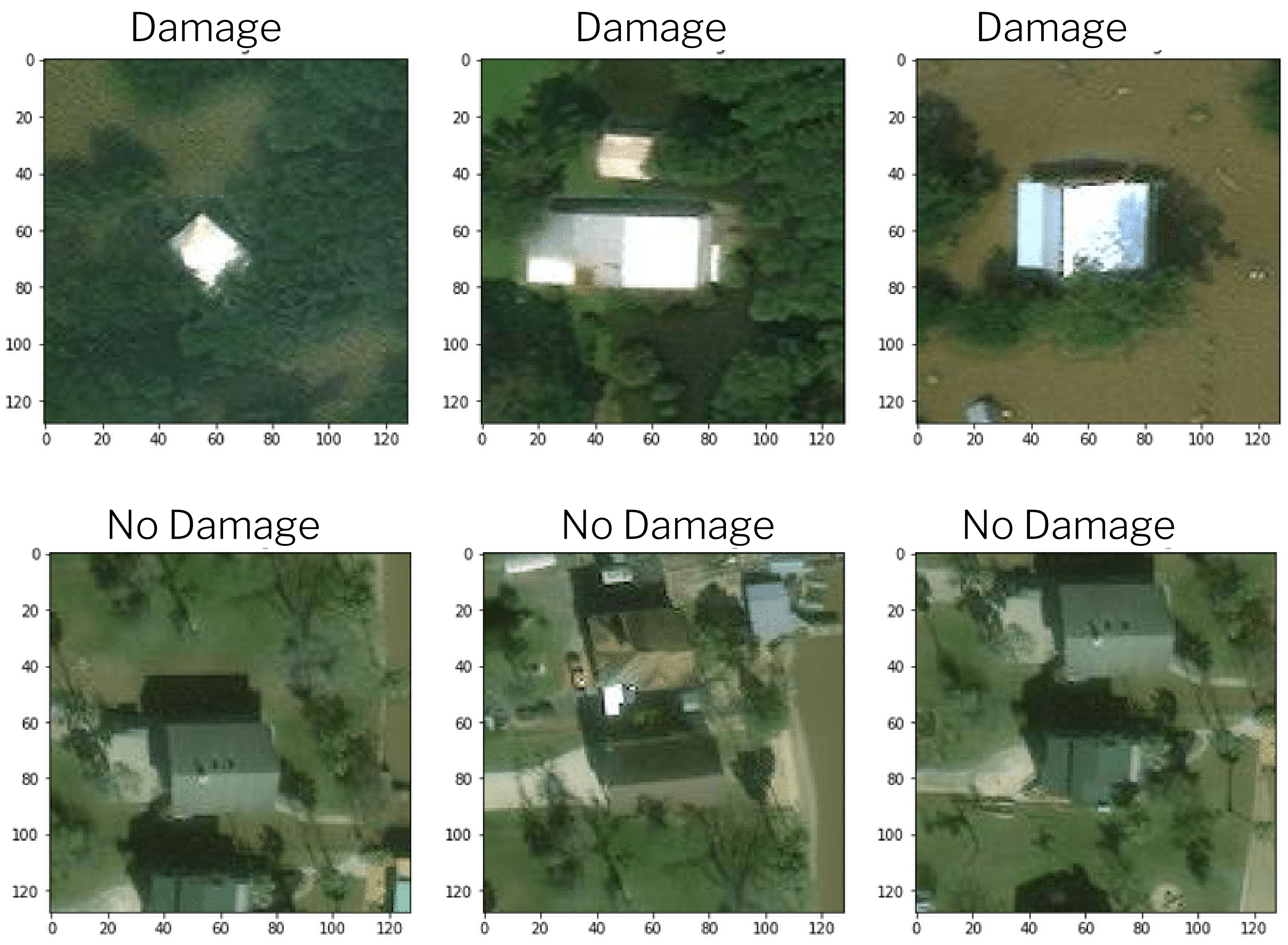
ACCURACY = 99.3%

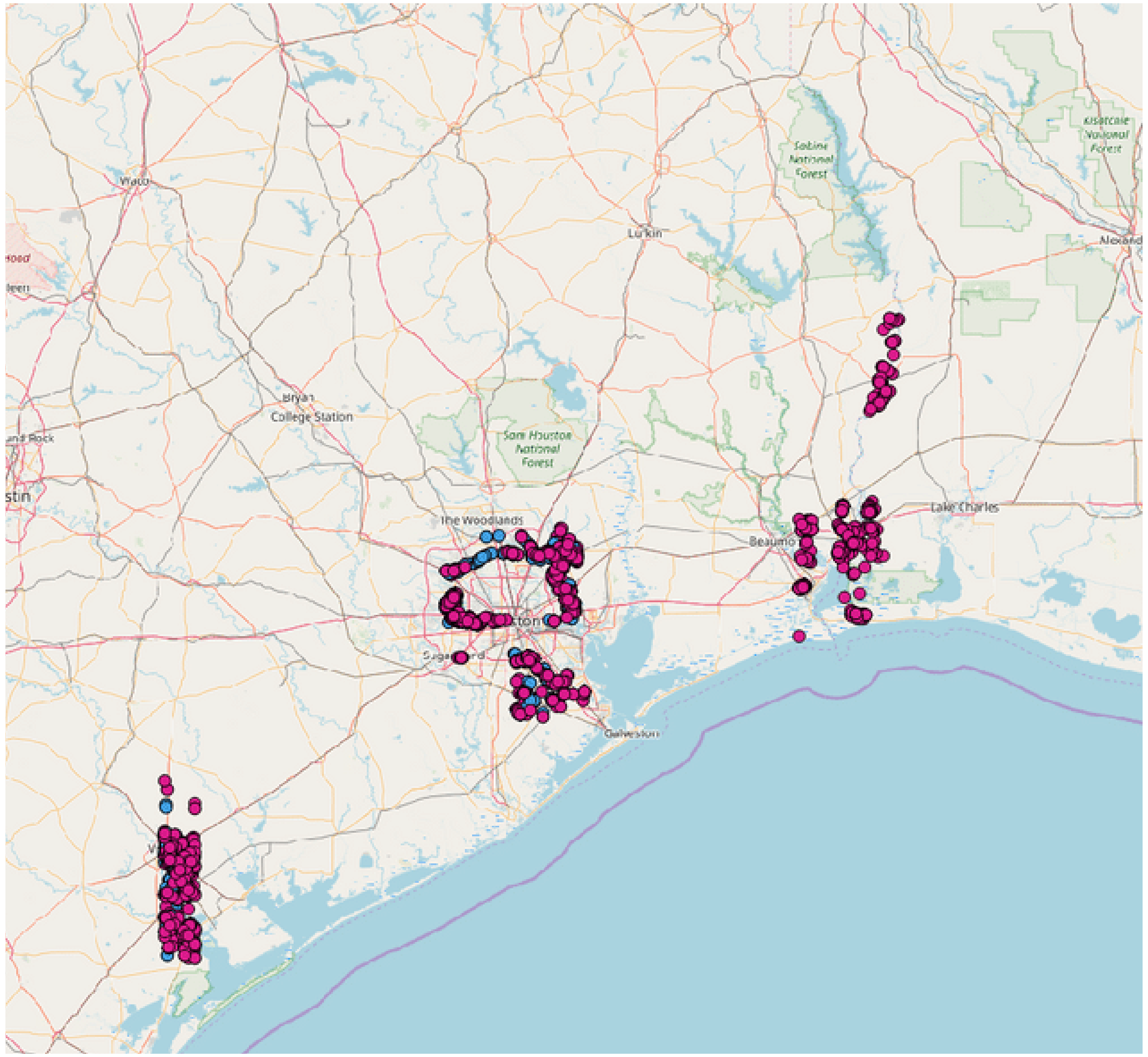
CONCLUSION

For our task, a designed pre-trained model shows high performance in classifying intact/damaged houses, with appropriate layers and data augmentation.

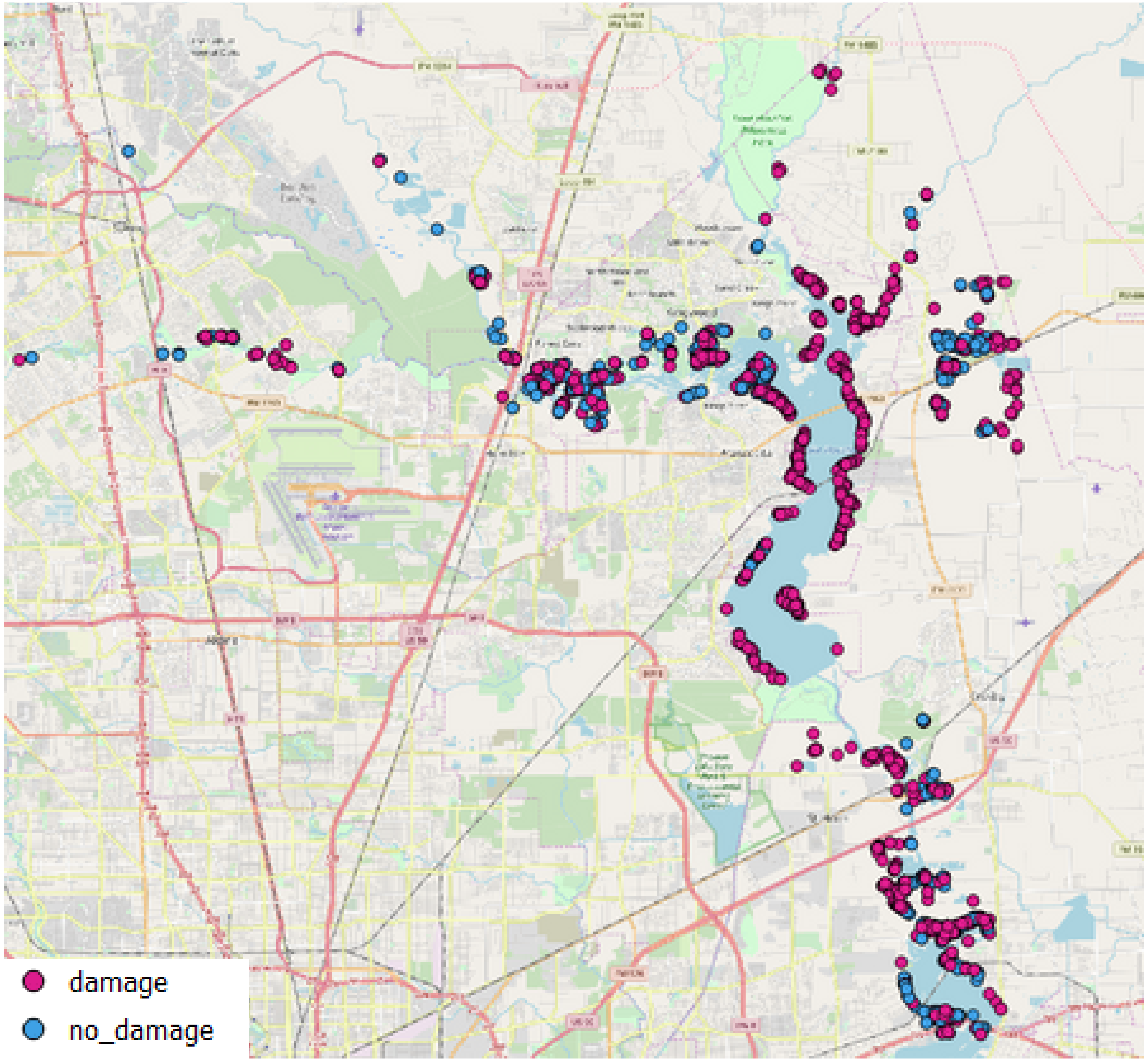
+99% of buildings were accurately classified in post-hurricane satellite imagery.

PREDICTIONS



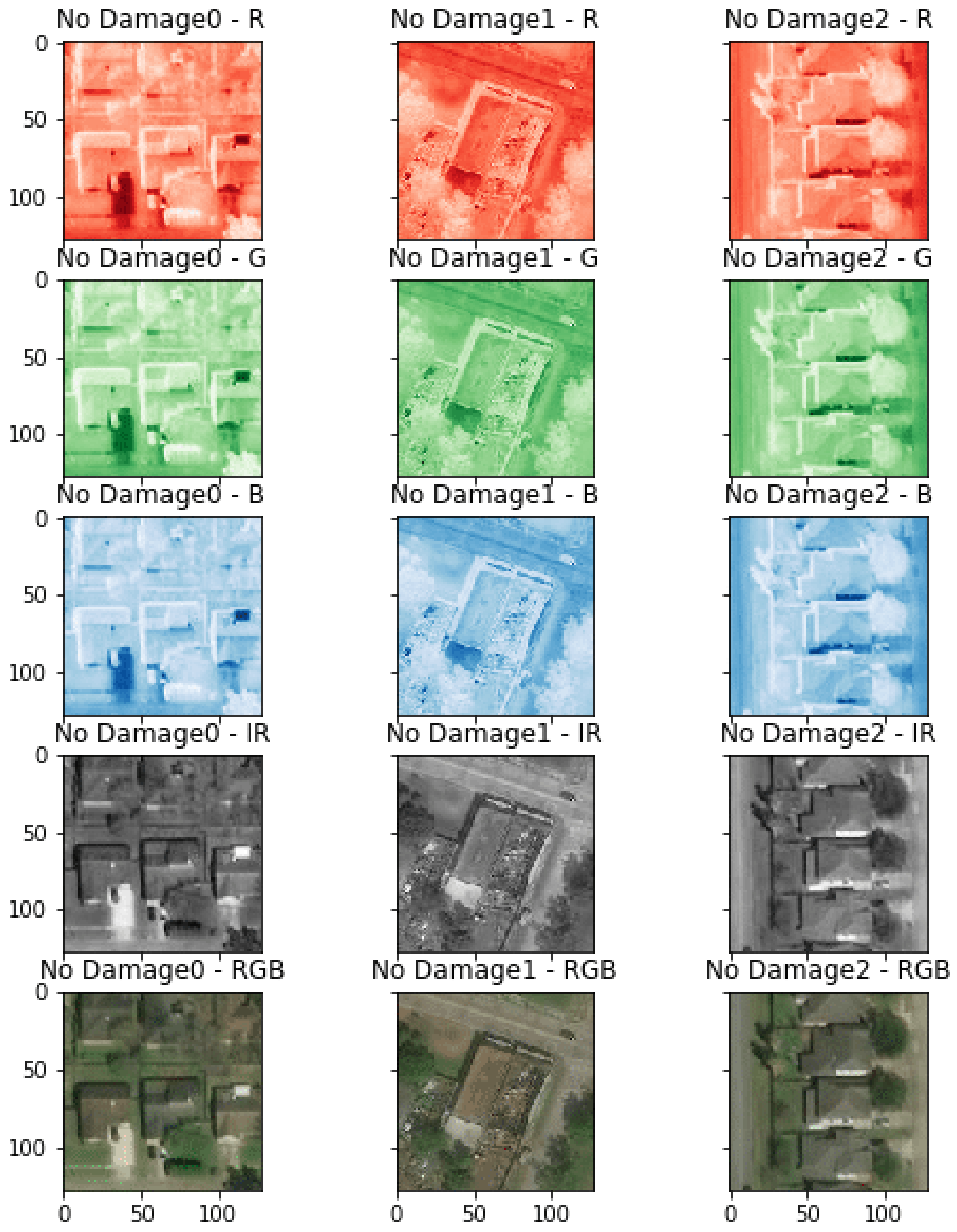


SATELLITE IMAGERY ALLOWS DAMAGE ASSESSMENT AND PRECISE LOCATION OF IMPACTED AREAS



FUTURE IMPROVEMENTS

- False negatives are due to confusion between water and smooth surfaces surrounding buildings, using solely RGB bands.
- For our task, **infrared channel** of satellite imagery might improve our neural networks.
- Bias = only american houses in training data. A solution might be integrating other countries buildings.



Sources

https://www.thebalance.com/hurricane-damage-economic-costs-4150369

https://ieee-dataport.org/open-access/detecting-damaged-buildings-post-hurricane-satellite-imagery-based-customized

https://medium.com/better-programming/why-you-should-use-convolutions-in-your-next-neural-net-using-tensorflow-37d347544454

https://www.kaggle.com/abhiksark/introduction-to-transfer-learning-cats-dogs/output

https://www.kaggle.com/pmigdal/transfer-learning-with-resnet-50-in-keras

Resources

Géron, A. (2017). Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems. " O'Reilly Media, Inc.".

He et al. (2015) Deep Residual Learning for Image Recognition

Simonyan & Zisserman (2014), Very Deep Convolutional Networks for Large Scale Image Recognition.

Bibliography

Cao, Q. D., & Choe, Y. (2018). Building Damage Annotation on Post-Hurricane Satellite Imagery Based on Convolutional Neural Networks. arXiv preprint arXiv:1807.01688.

Merci !

Joos, Marwan, Radia & Jean-Marc

Marjorie, Oriane, Morgane, Daygina, Laura, Lauma, Sandrine, Julie T., Julie G., Aline, Nancy, Stressie, Laure, Flora, Elodie <3

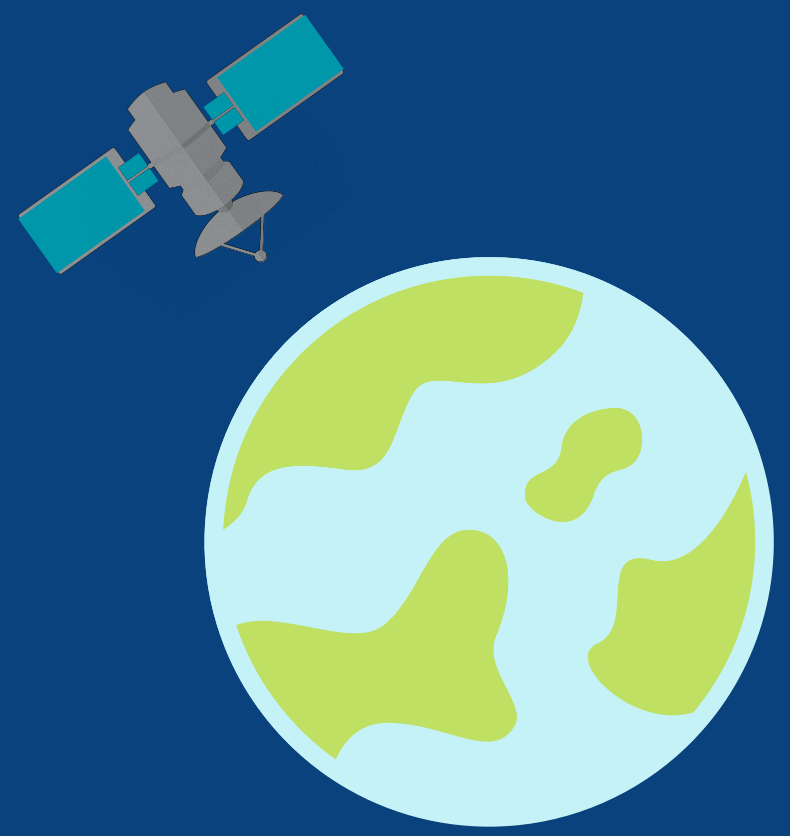
Appendix

DATASET

labeled buildings as flooded or not (50%/50%)

10 000 images **2000 images**

TRAIN VALIDATION



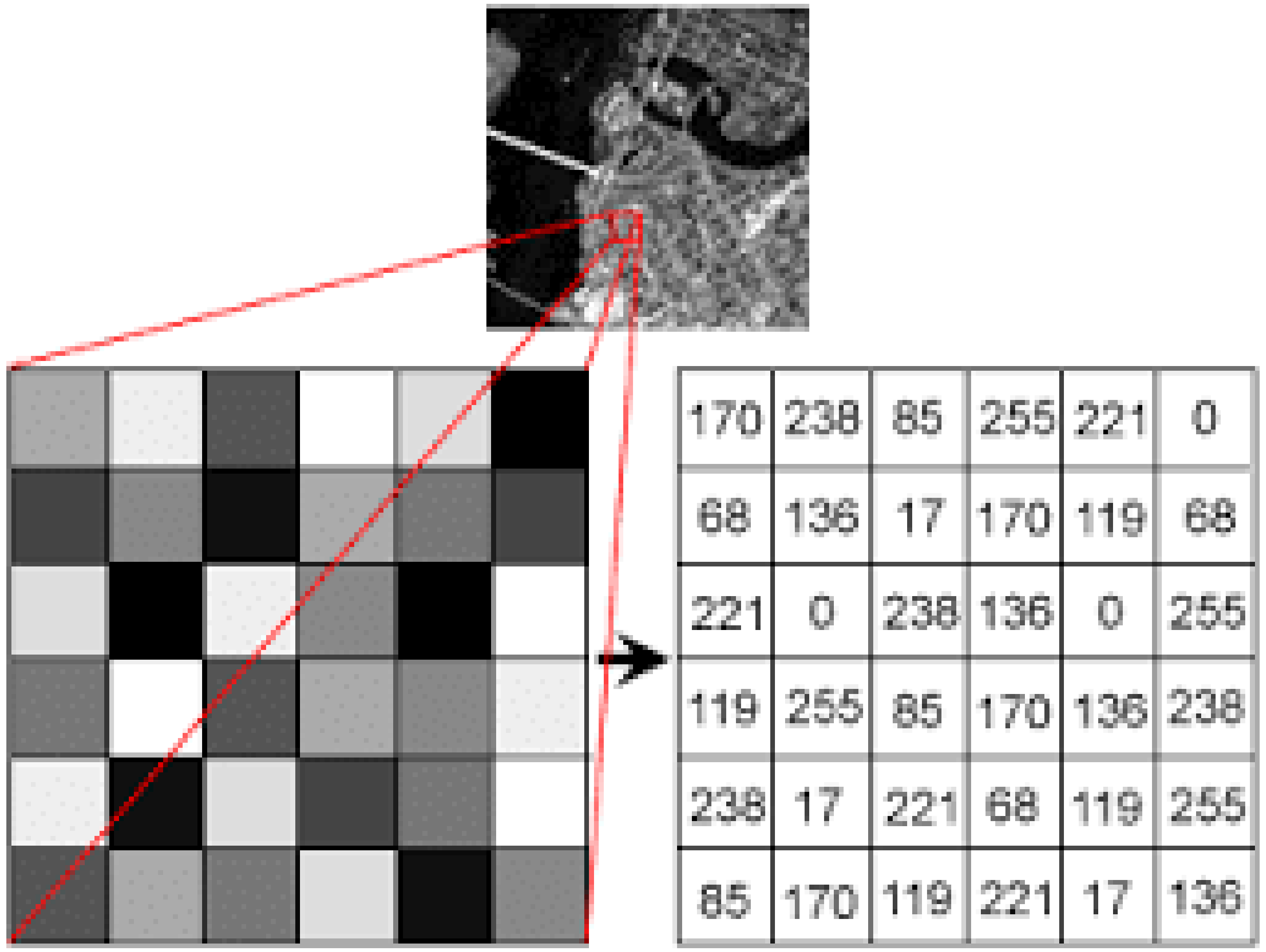
UNKNOWN DATA

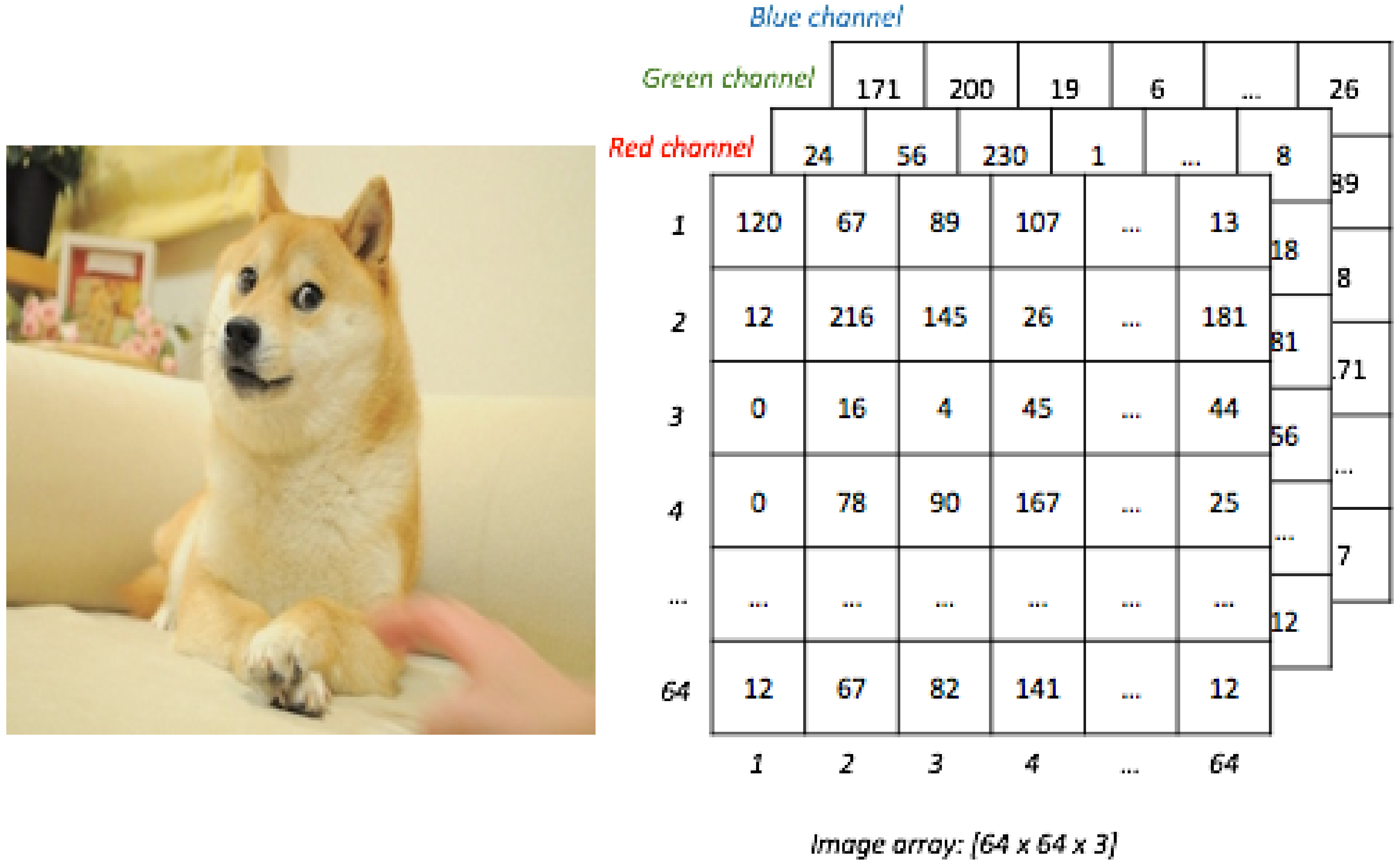
buildings to be classified (50%/50%)

2000 images

TEST

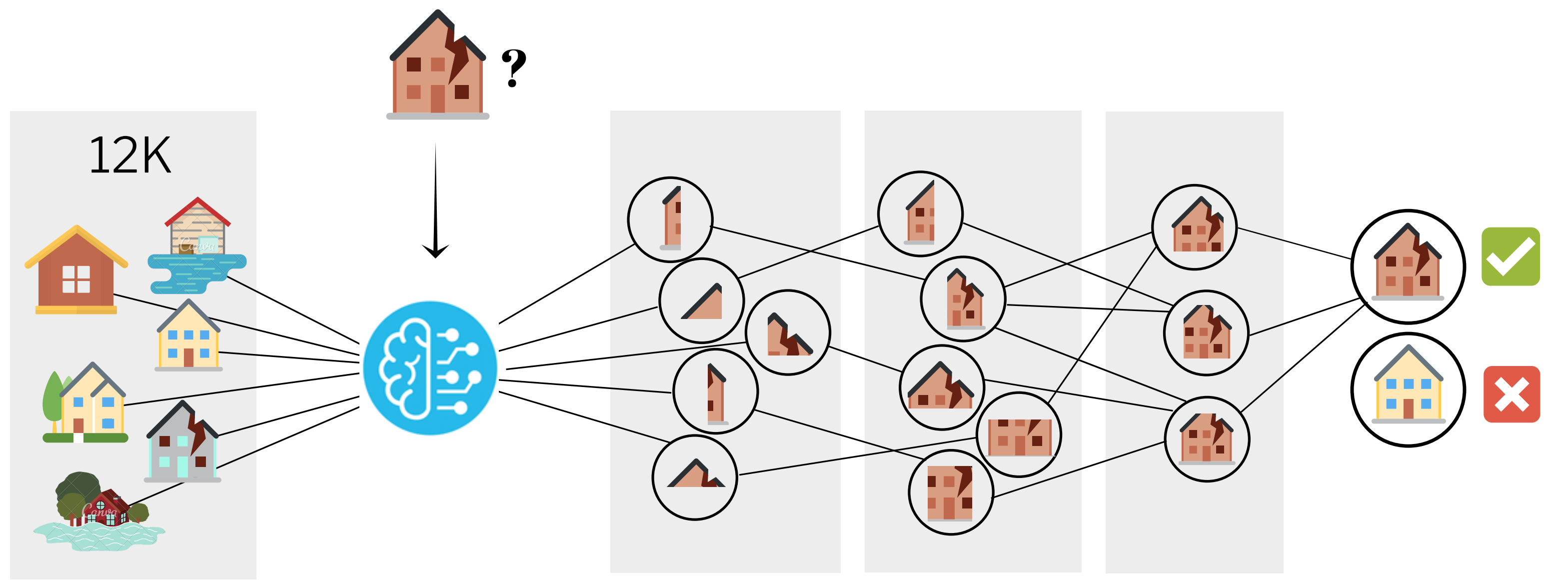
WHAT THE COMPUTER SEES





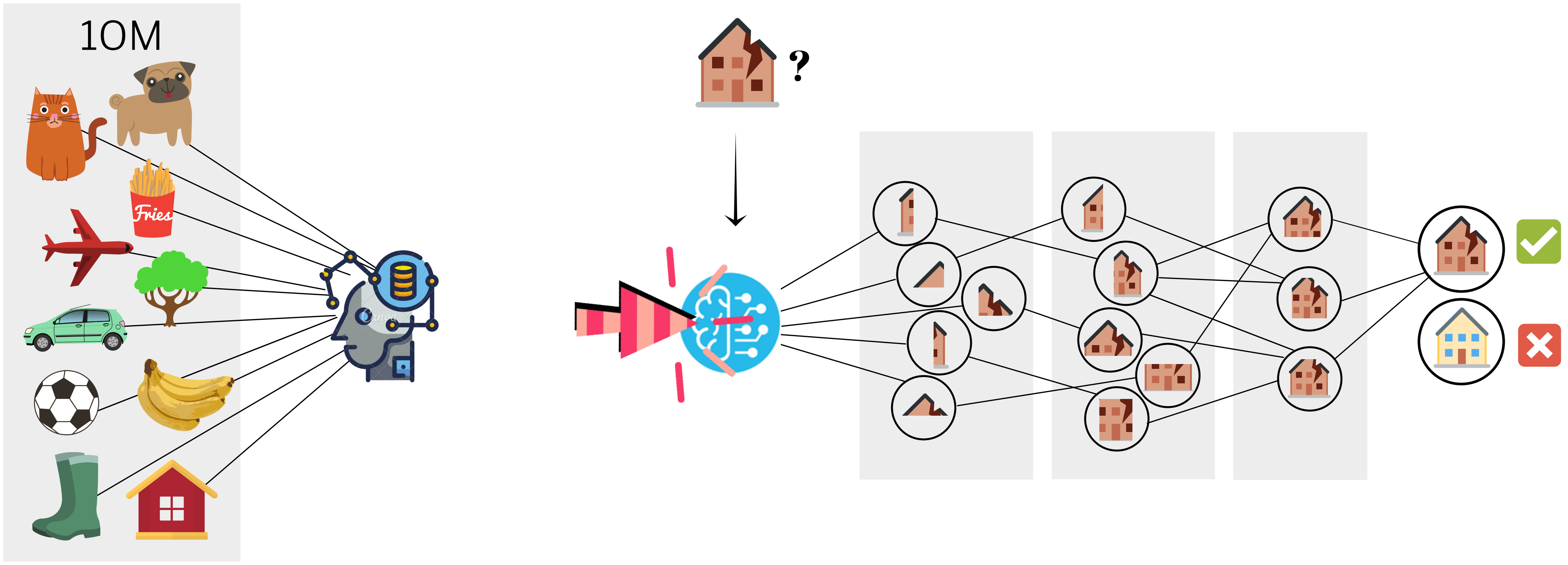
MODEL TESTING APPROACH

- **BUILDING OUR OWN MODEL**



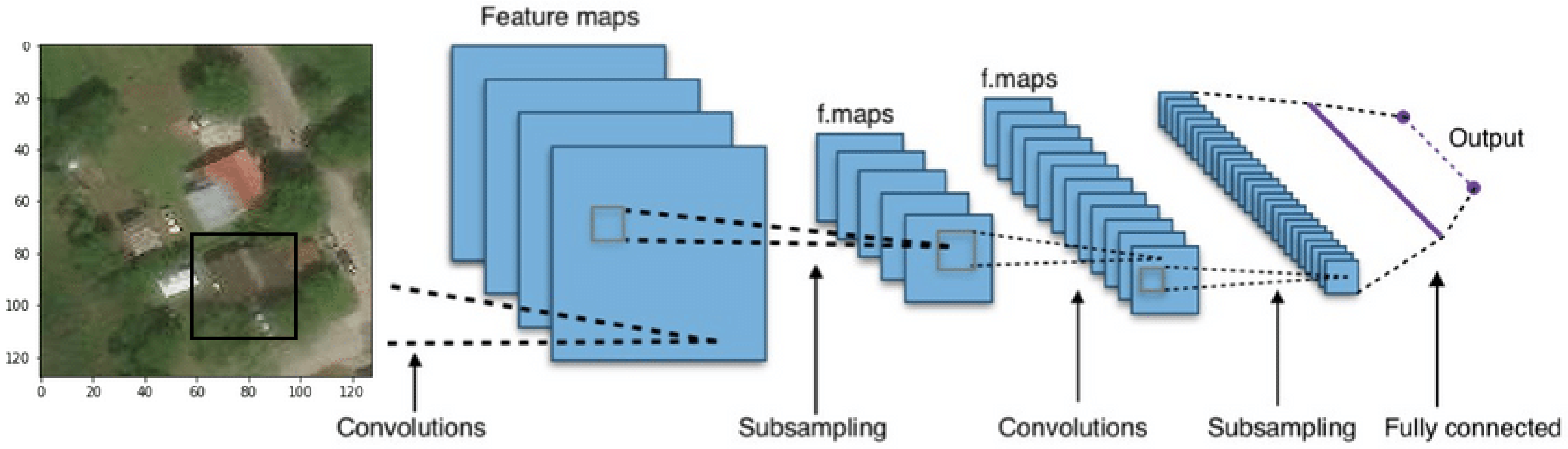
- **USING PRE-TRAINED MODELS IN OUR DESIGNED MODEL**

  Also called Transfer learning: reusing the already acquired knowledge.



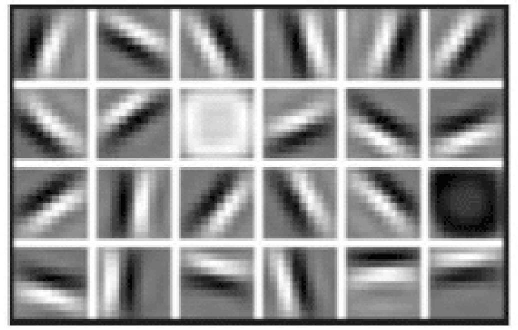
USING CONVOLUTIONAL NEURAL NETWORKS

Able to develop an internal representation of a 2-d image.

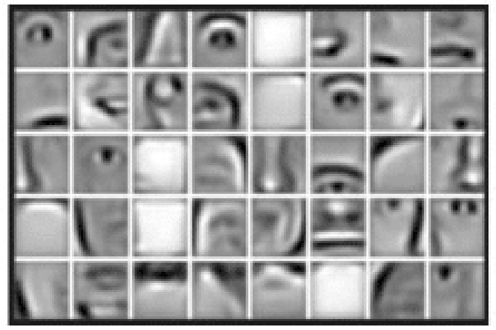


source: wikipedia

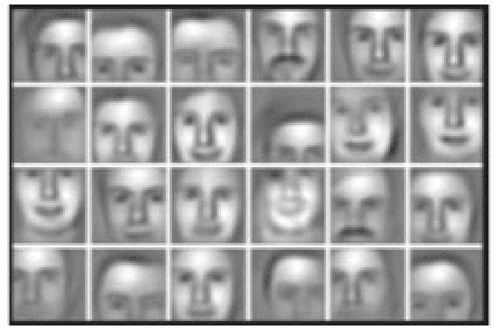
Convolutions use filters to highlight features: vertical or horizontal lines, curves...



First Layer Representation



Second Layer Representation

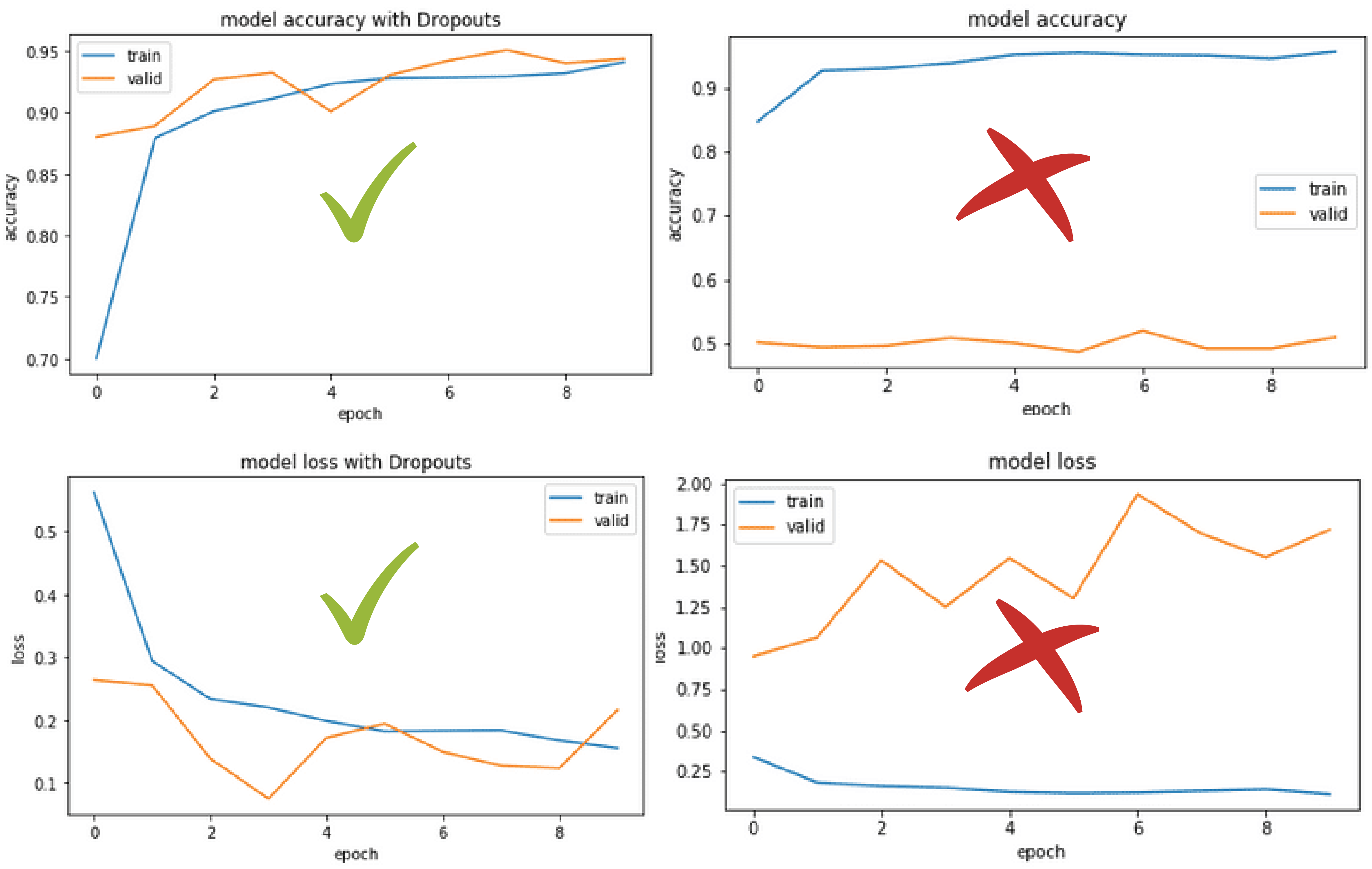


Third Layer Representation

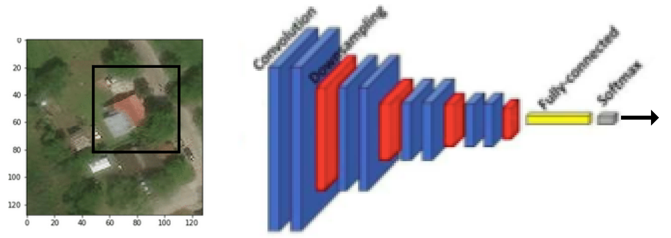
EVALUATION OF PERFORMANCE

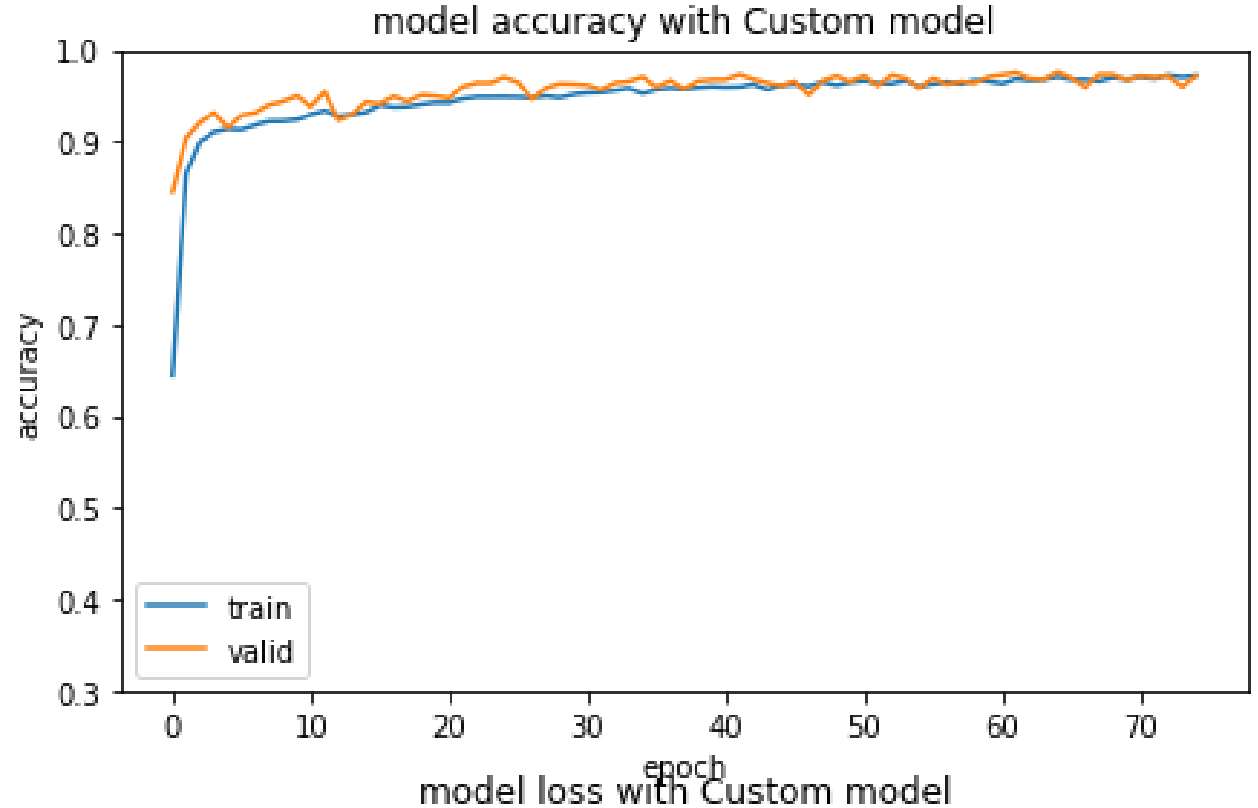
The metrics: accuracy and loss

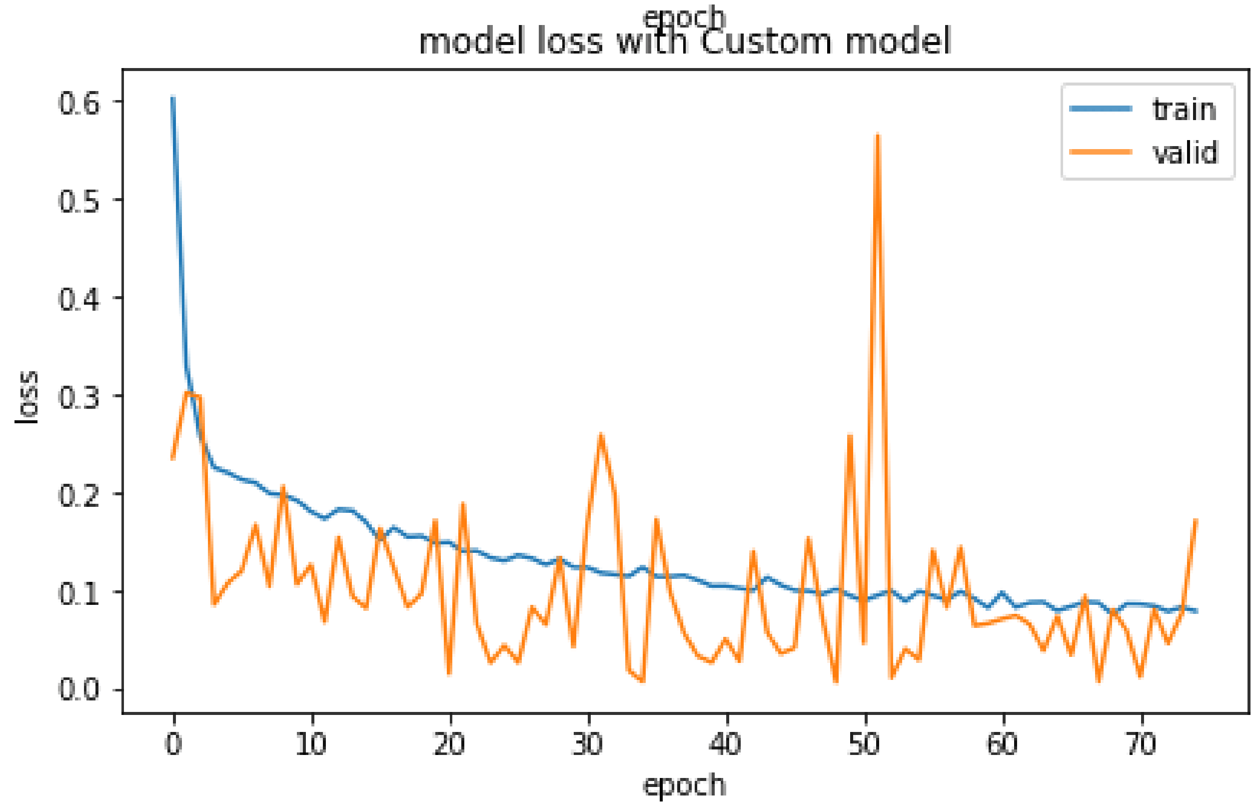
If model is overfitting, graph shows great performance on training data and poor performance on test data.



DESIGNED CNN





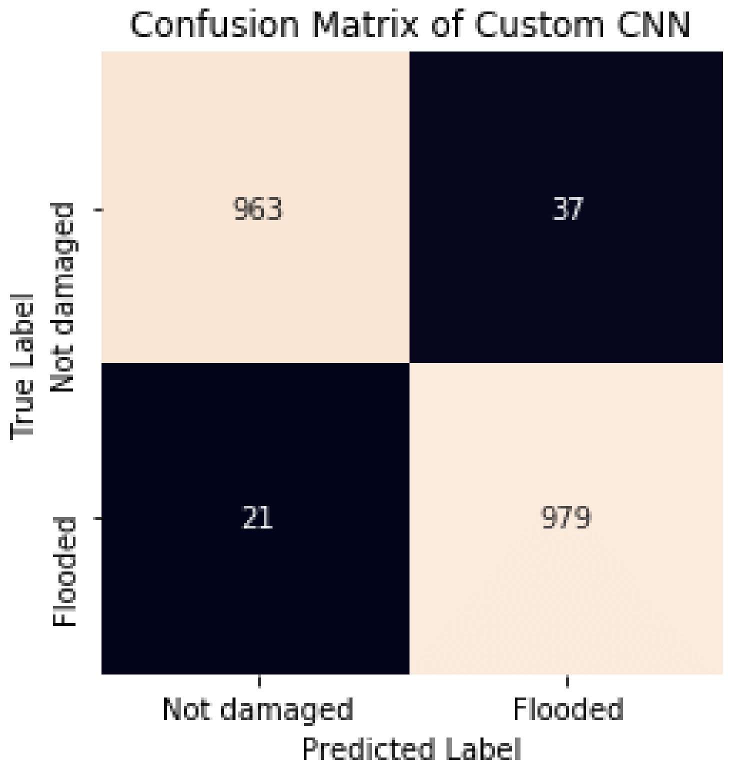


MODEL CUSTOM AFTER 75 EPOCHS

Validation accuracy = 97.5%

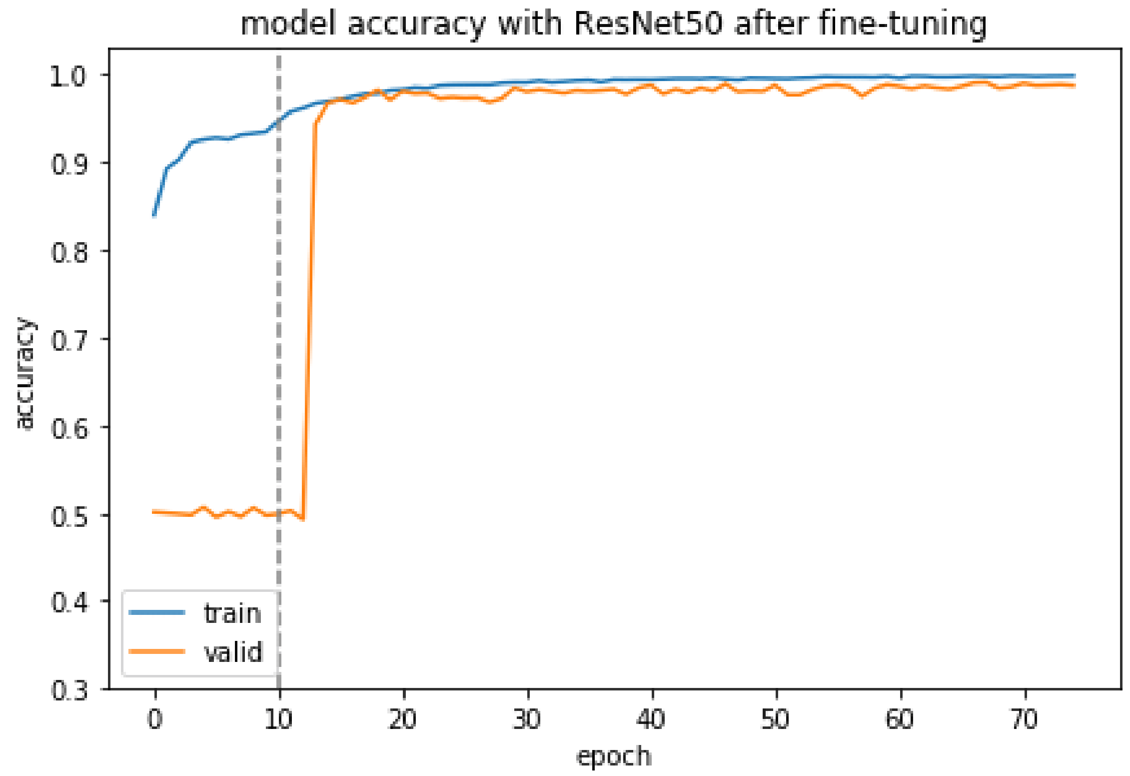
Test accuracy = 97.1%

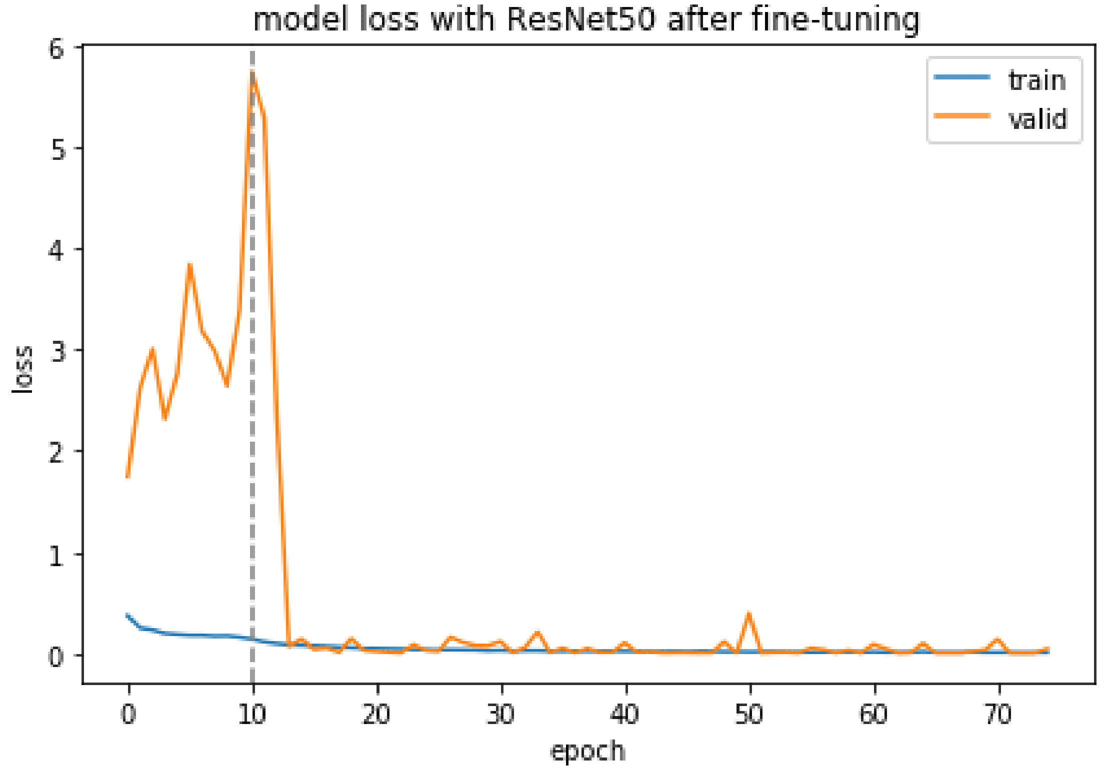
Training loss is very optimized but validation loss is a bit higher which indicates overfitting.



2 CUSTOM RESNET50

He et al. (2015)

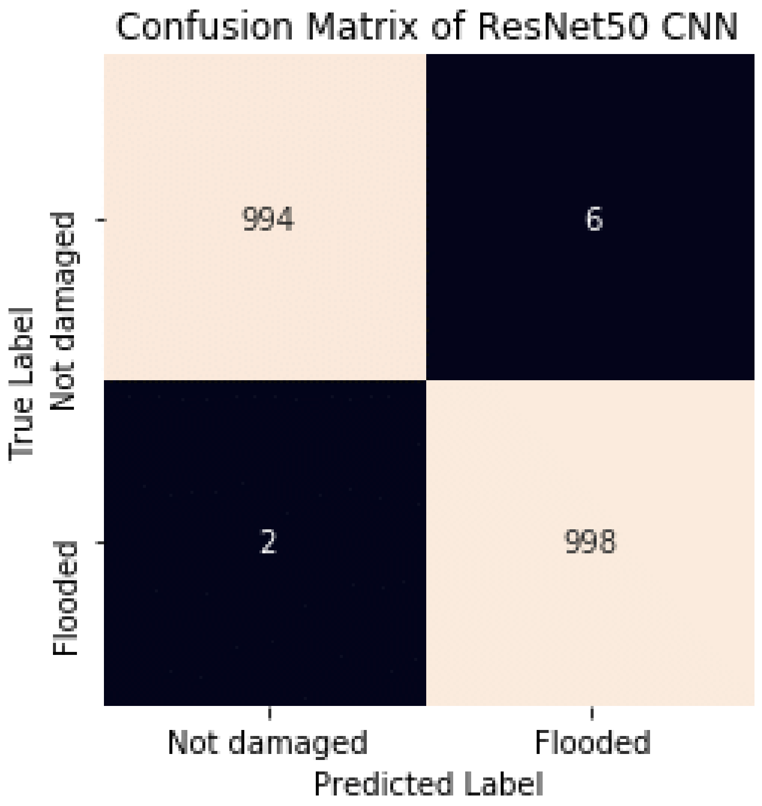




MODEL CUSTOM AFTER 75 EPOCHS

Validation accuracy = 99.0%

Test accuracy = 99.6%



Our Custom ResNet50 outperforms other models with:

1. higher accuracy and lower loss on validation data, and
2. better predictions on test data.