Hybrid Networks

Improving Deep Learning Networks via Integrating two views of Images

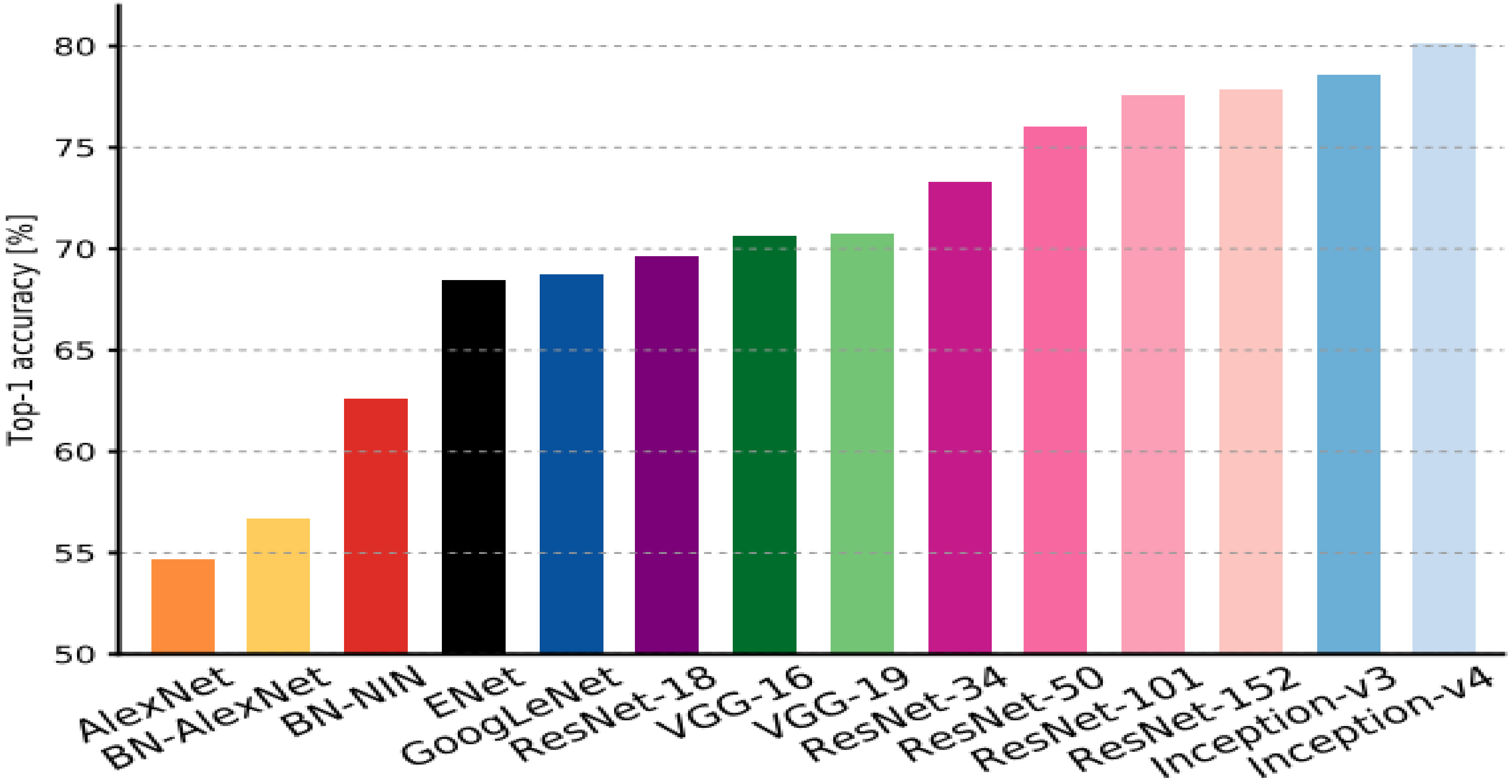
Sunny Verma[1,2], Chen Wang[2], Liming Zhu[2],and Wei Liu[1]

1: Advanced Analytics Institute, University of Technology, Sydney, Australia

2: Architecture and Analytics Platforms, CSIRO, Data61

Motivations

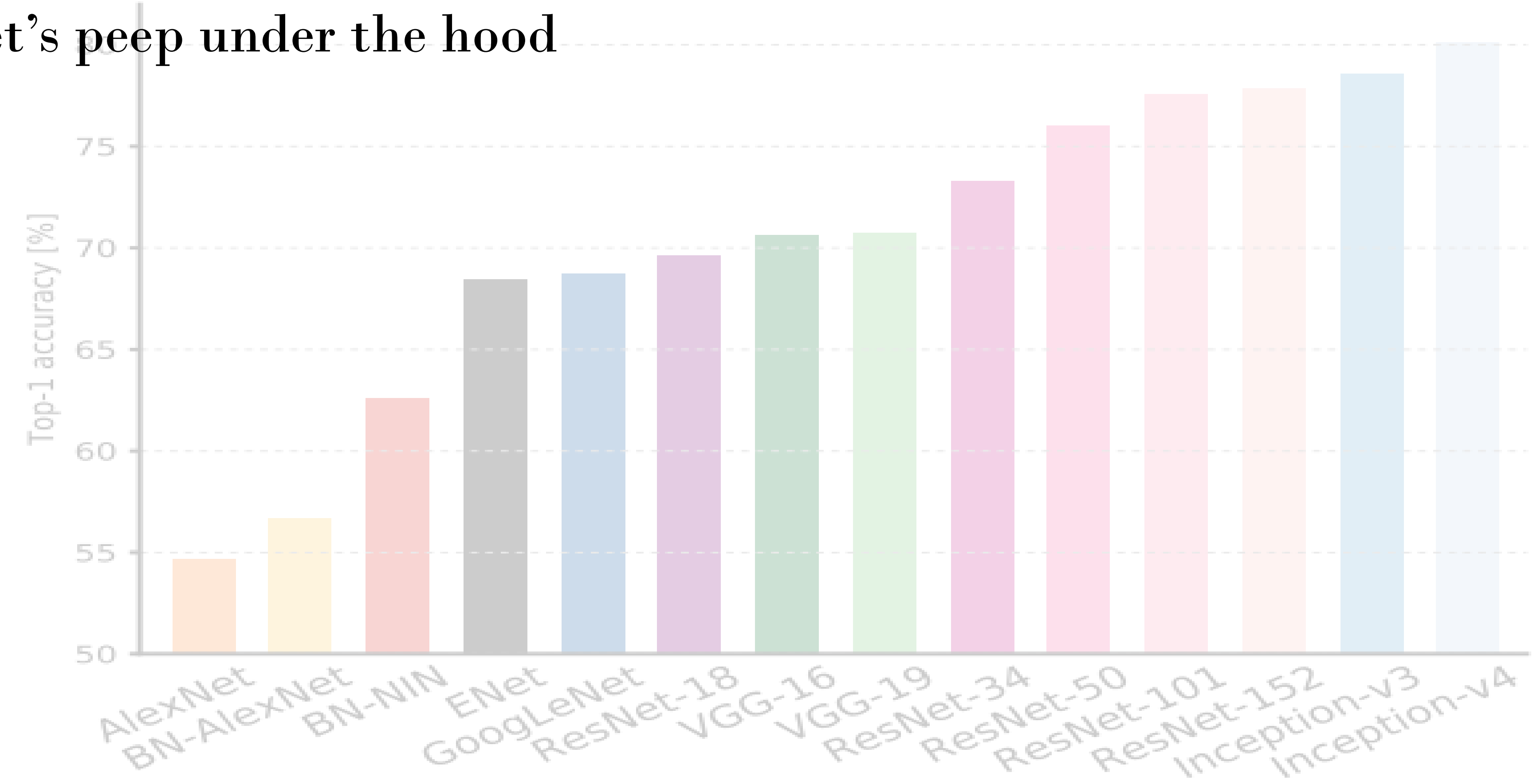
Deep Models have advanced image recognition field substantially



Motivations

Deep Models have advanced image recognition field substantially

- Let’s peep under the hood

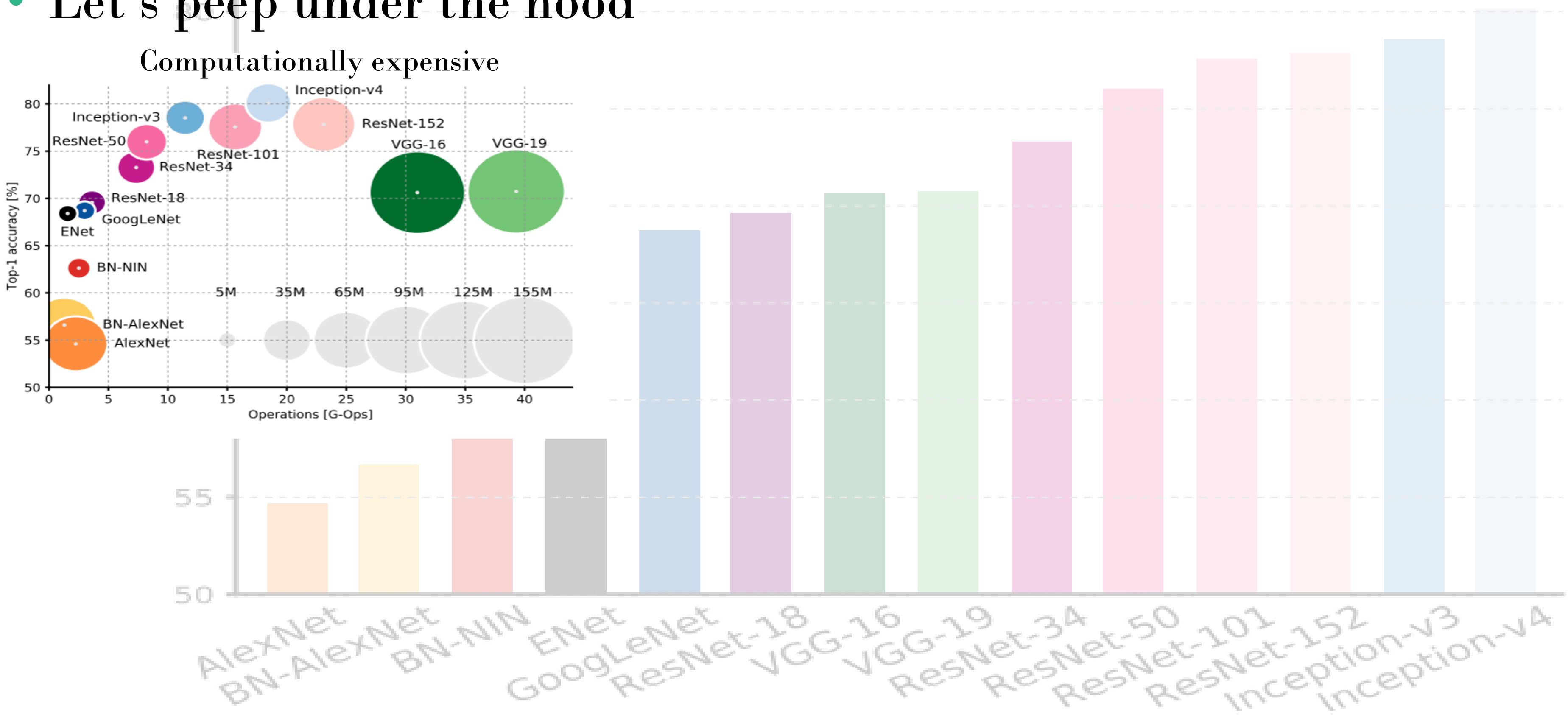


Motivations

Deep Models have advanced image recognition field substantially

- Let's peep under the hood

Computationally expensive

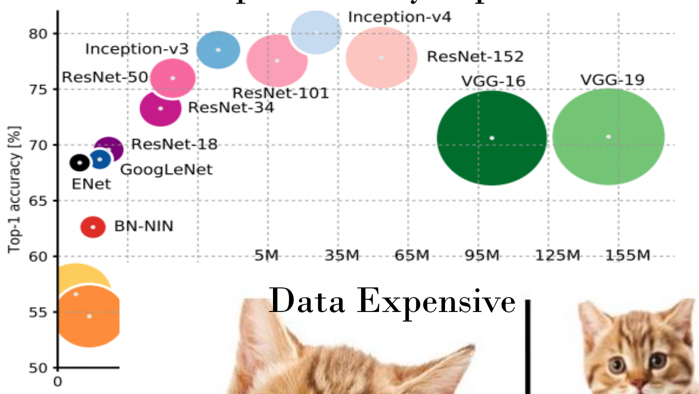


Motivations

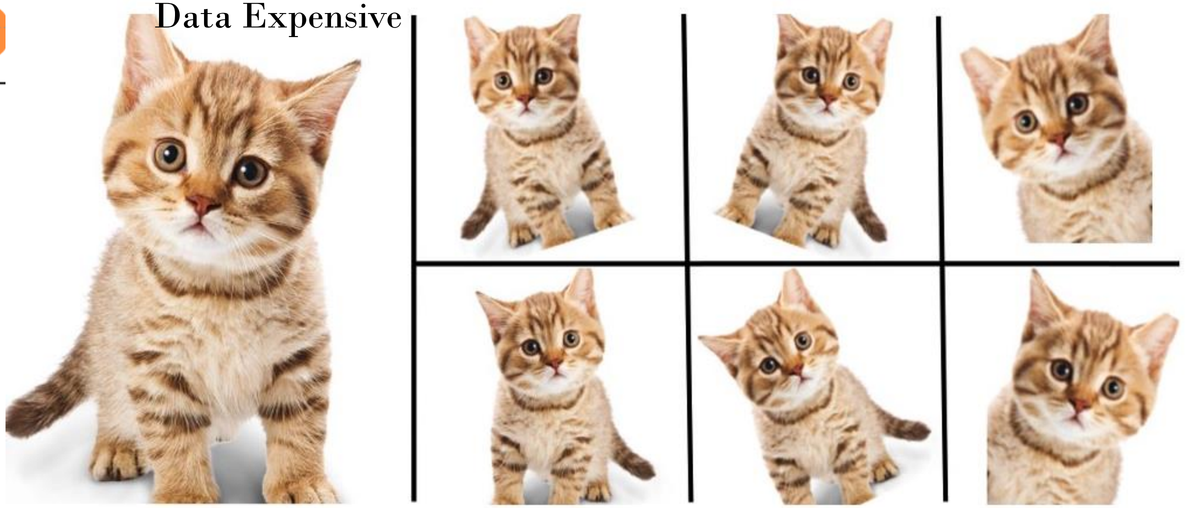
Deep Models have advanced image recognition field substantially

- Let’s peep under the hood

Computationally expensive



Data Expensive

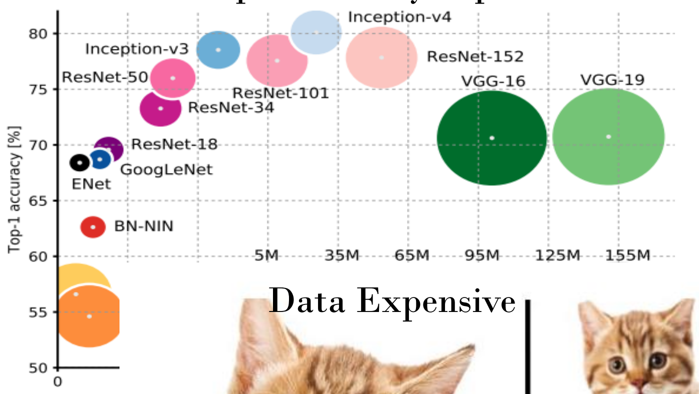


Motivations

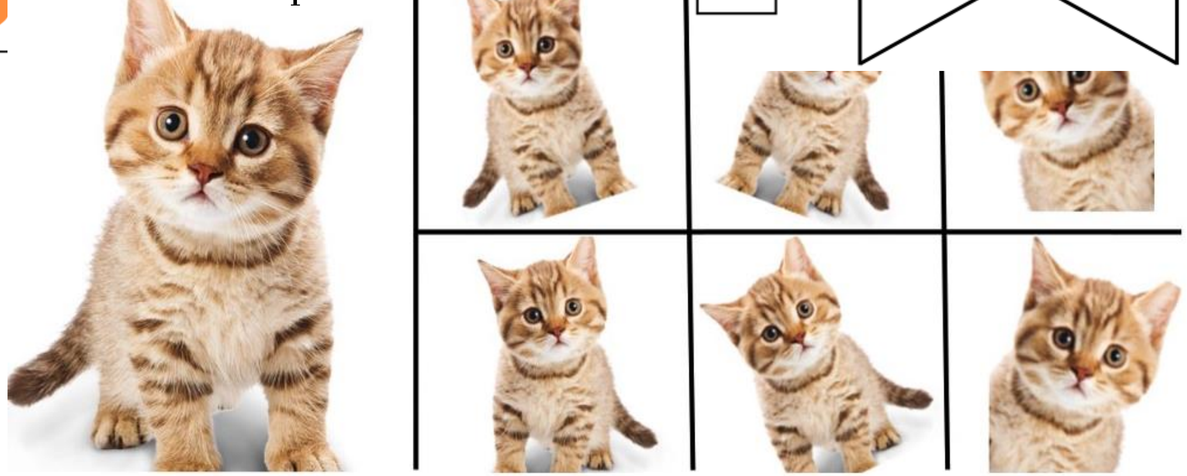
Deep Models have advanced image recognition field substantially

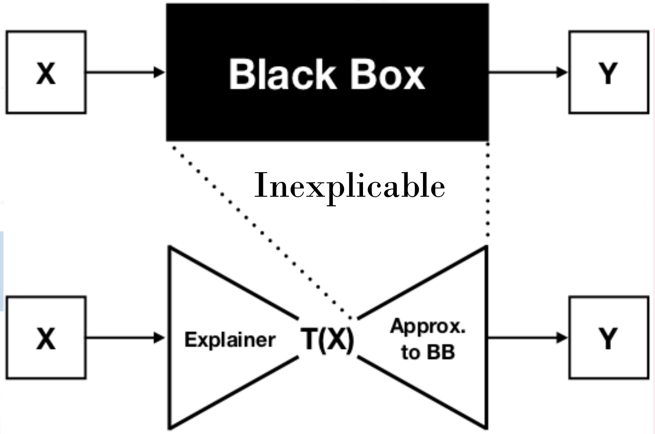
- Let’s peep under the hood

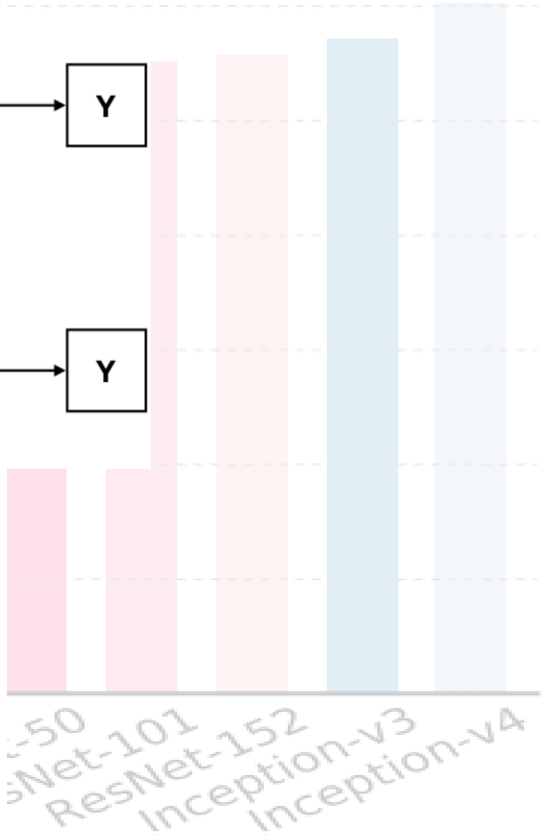
Computationally expensive



Data Expensive





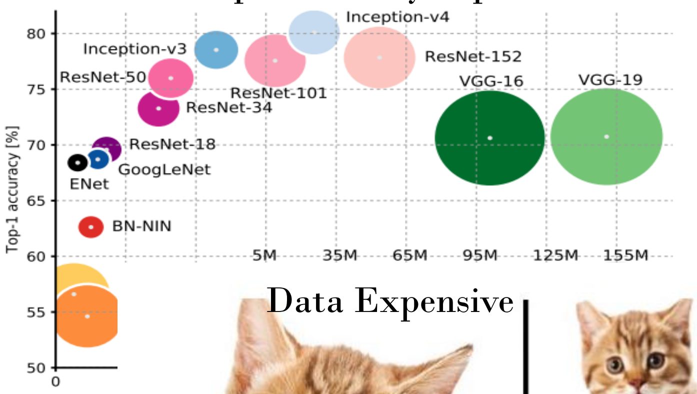


Motivations

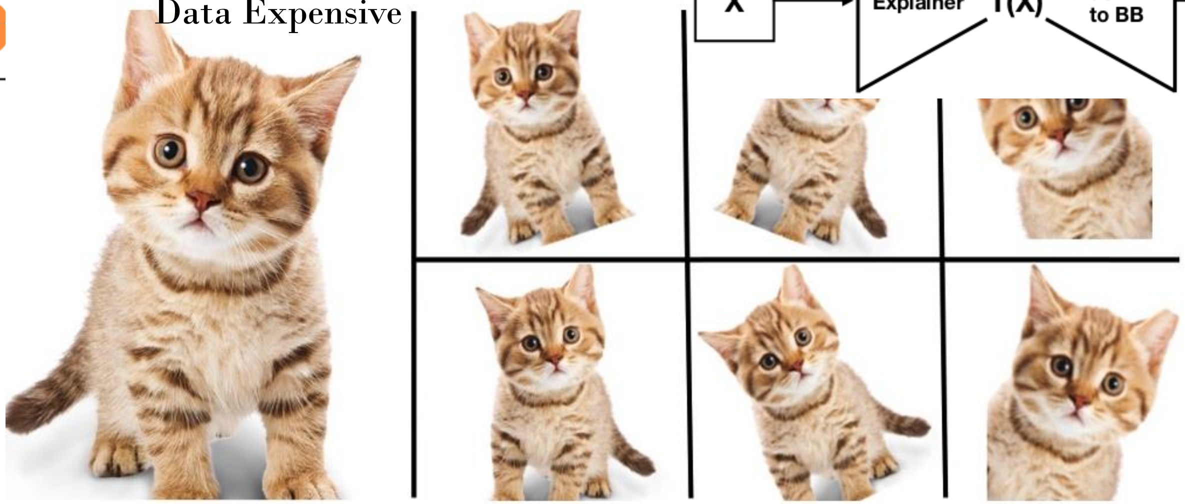
Deep Models have advanced image recognition field substantially

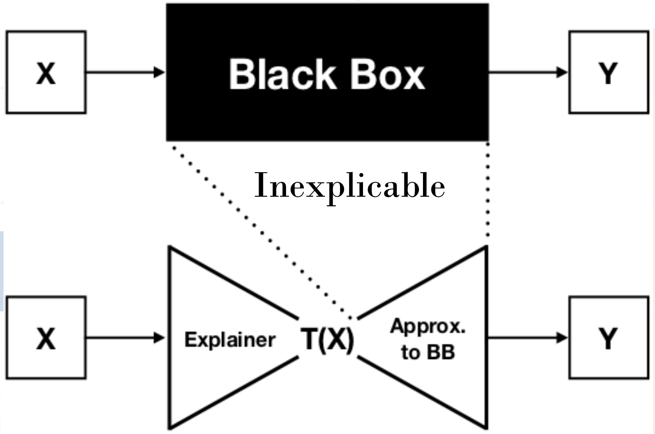
- Let's peep under the hood

Computationally expensive

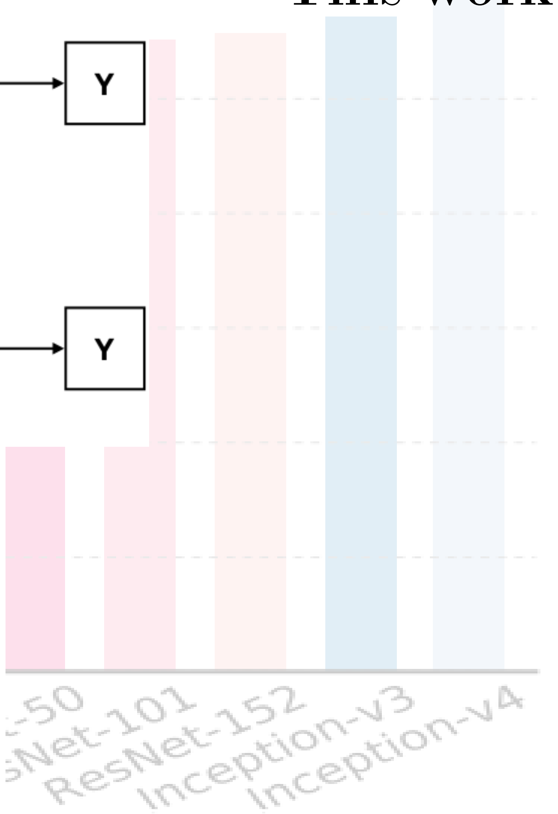


Data Expensive





- This work

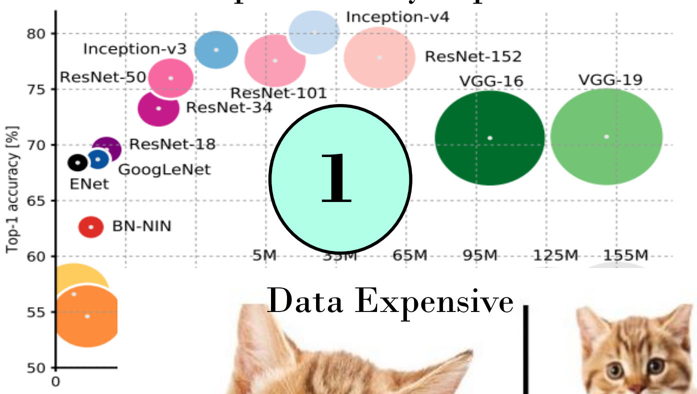


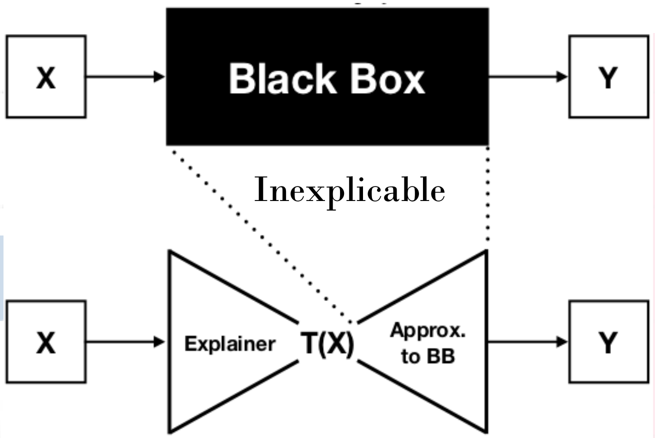
Motivations

Deep Models have advanced image recognition field substantially

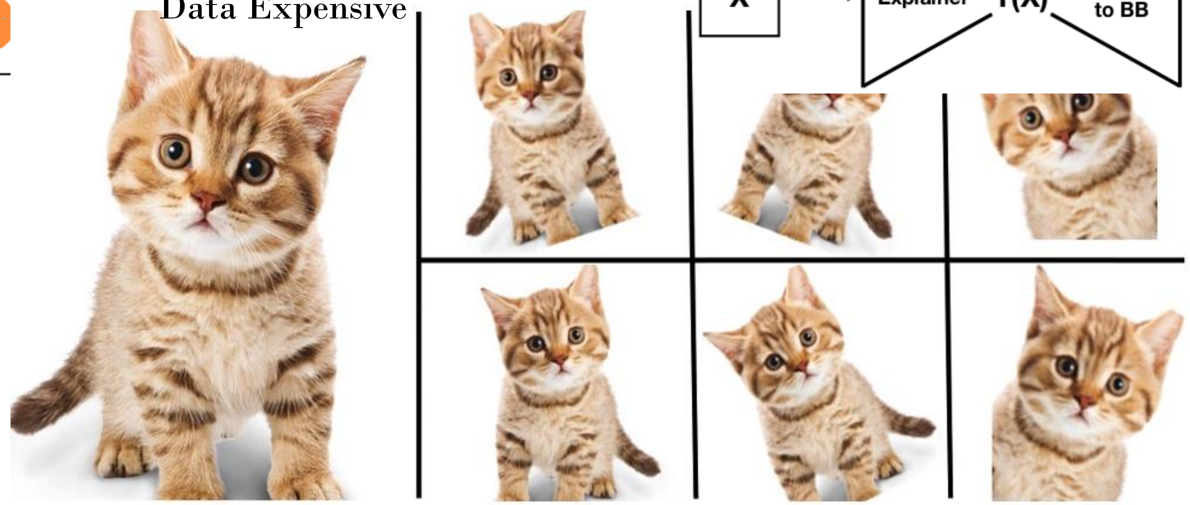
- Let’s peep under the hood

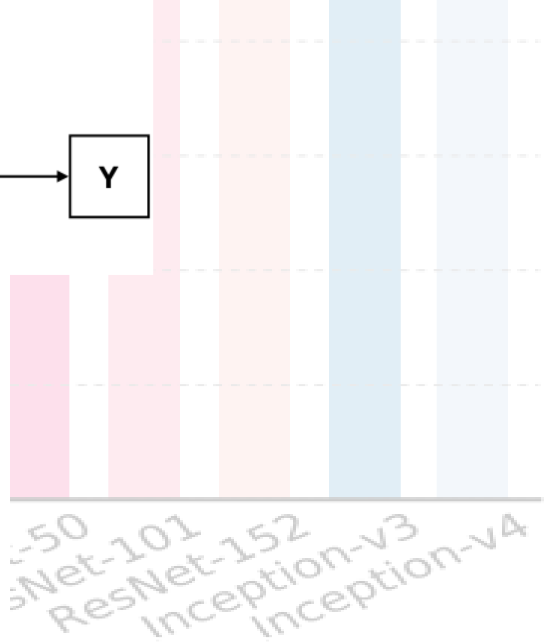
Computationally expensive





Data Expensive





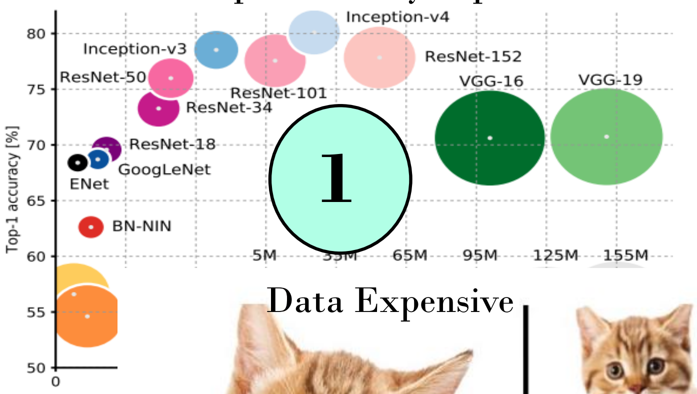
- This work
  1. Requires significantly less computational resources.

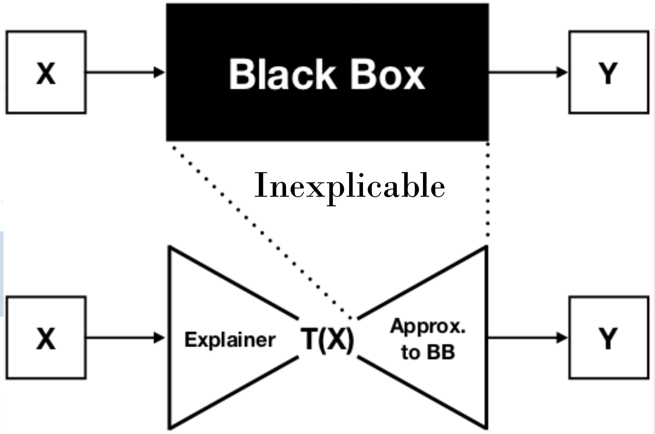
Motivations

Deep Models have advanced image recognition field substantially

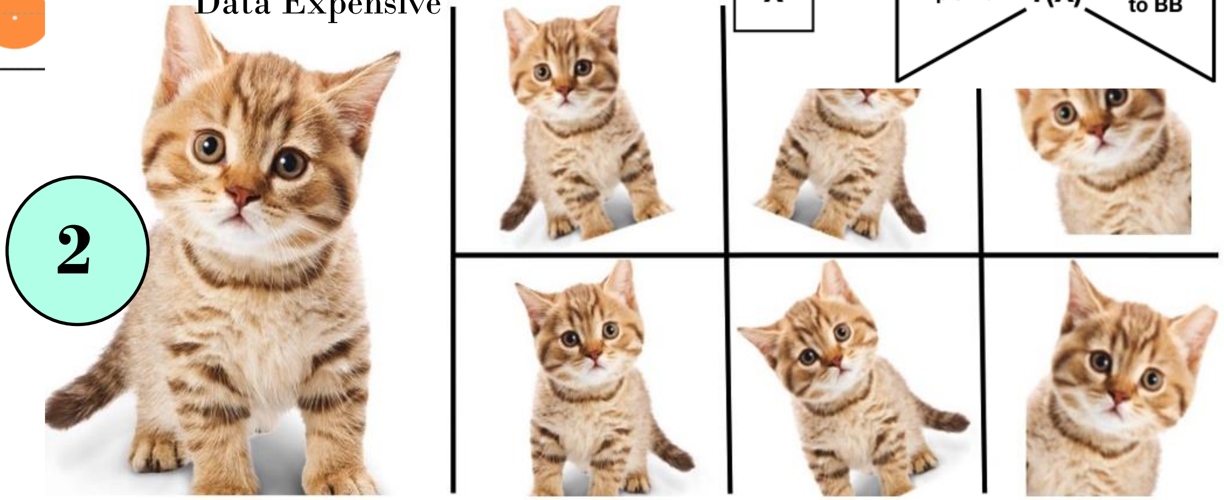
- Let's peep under the hood

Computationally expensive





Data Expensive



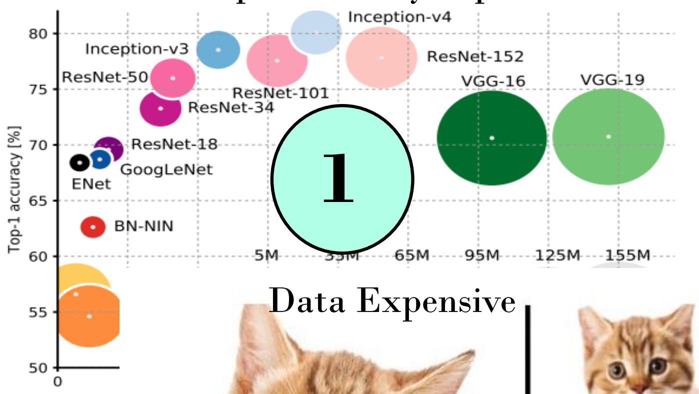
- This work
  1. Requires significantly less computational resources.
  2. Can work with available amount of data.

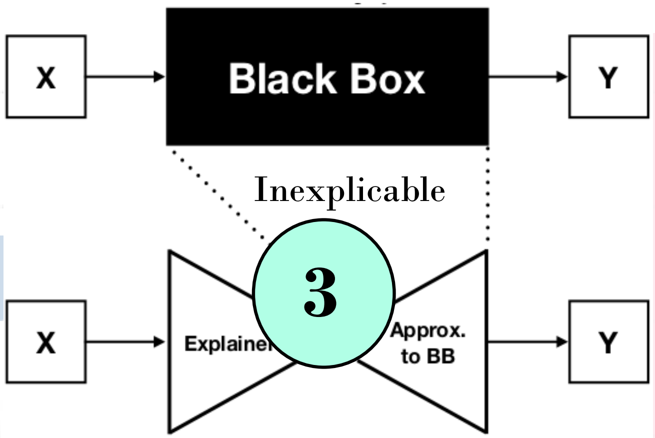
Motivations

Deep Models have advanced image recognition field substantially

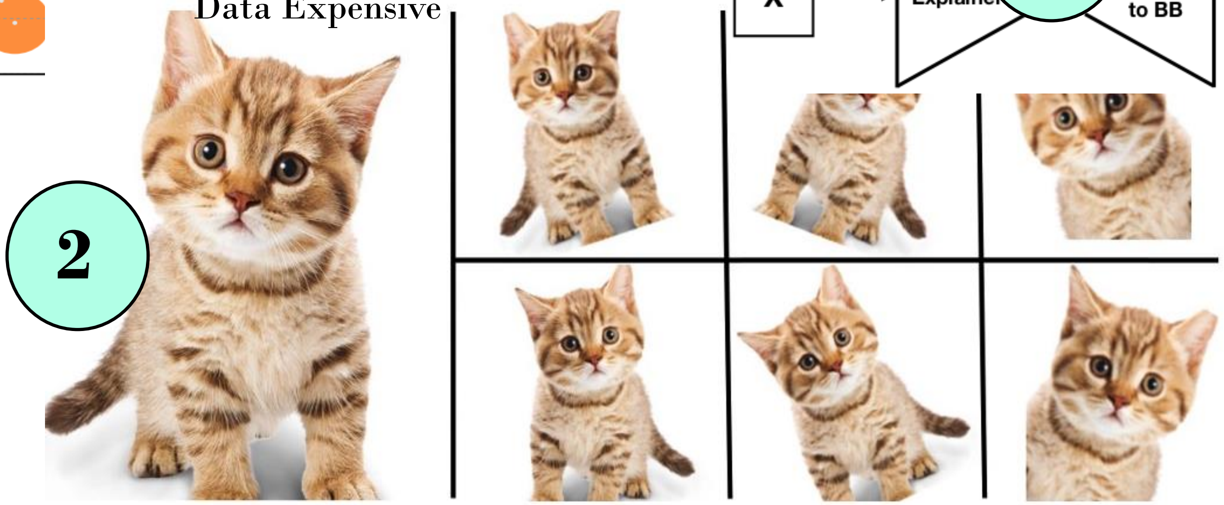
- Let's peep under the hood

Computationally expensive





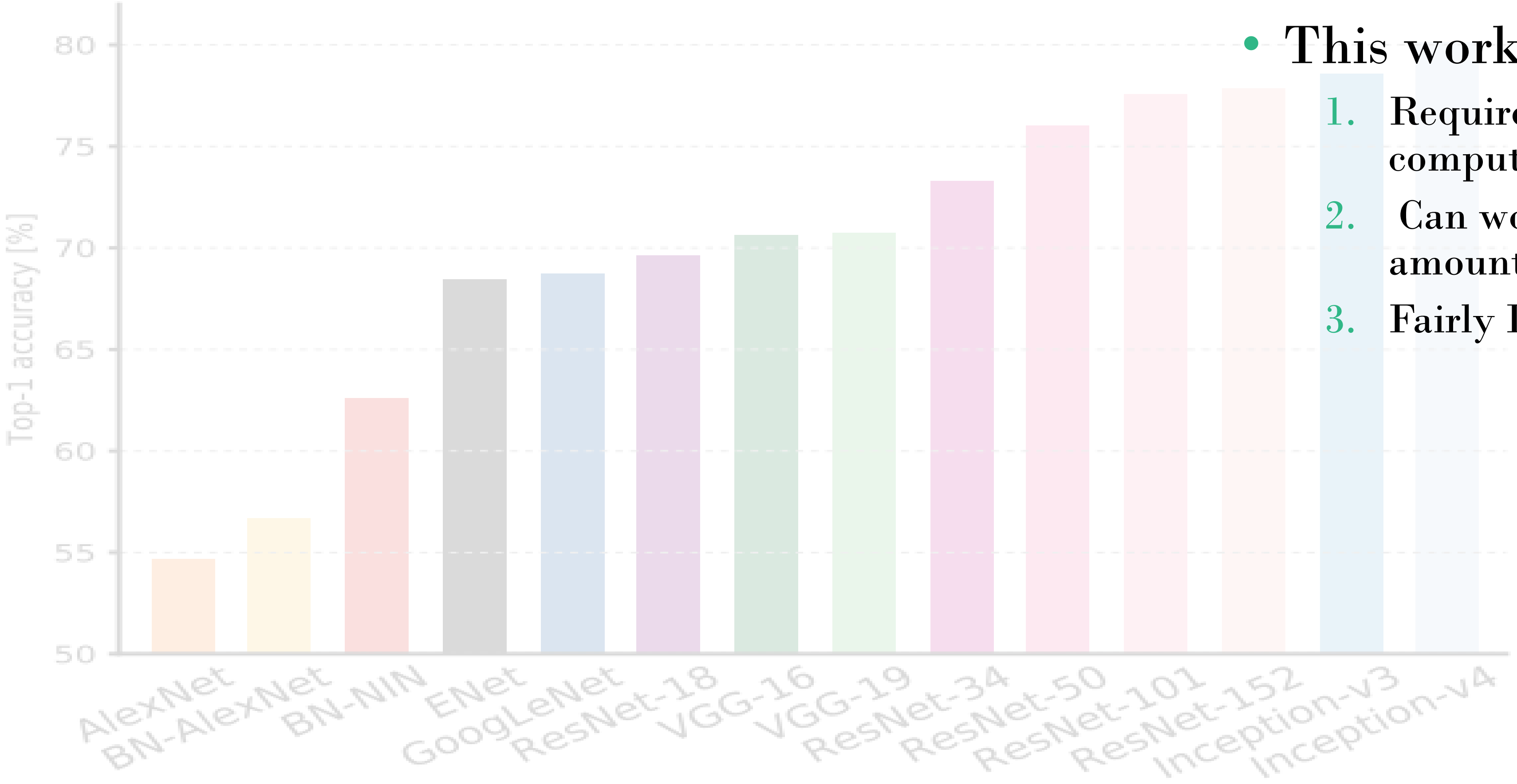
Data Expensive



- This work
  1. Requires significantly less computational resources.
  2. Can work with available amount of data.
  3. Fairly Interpretable

Challenges

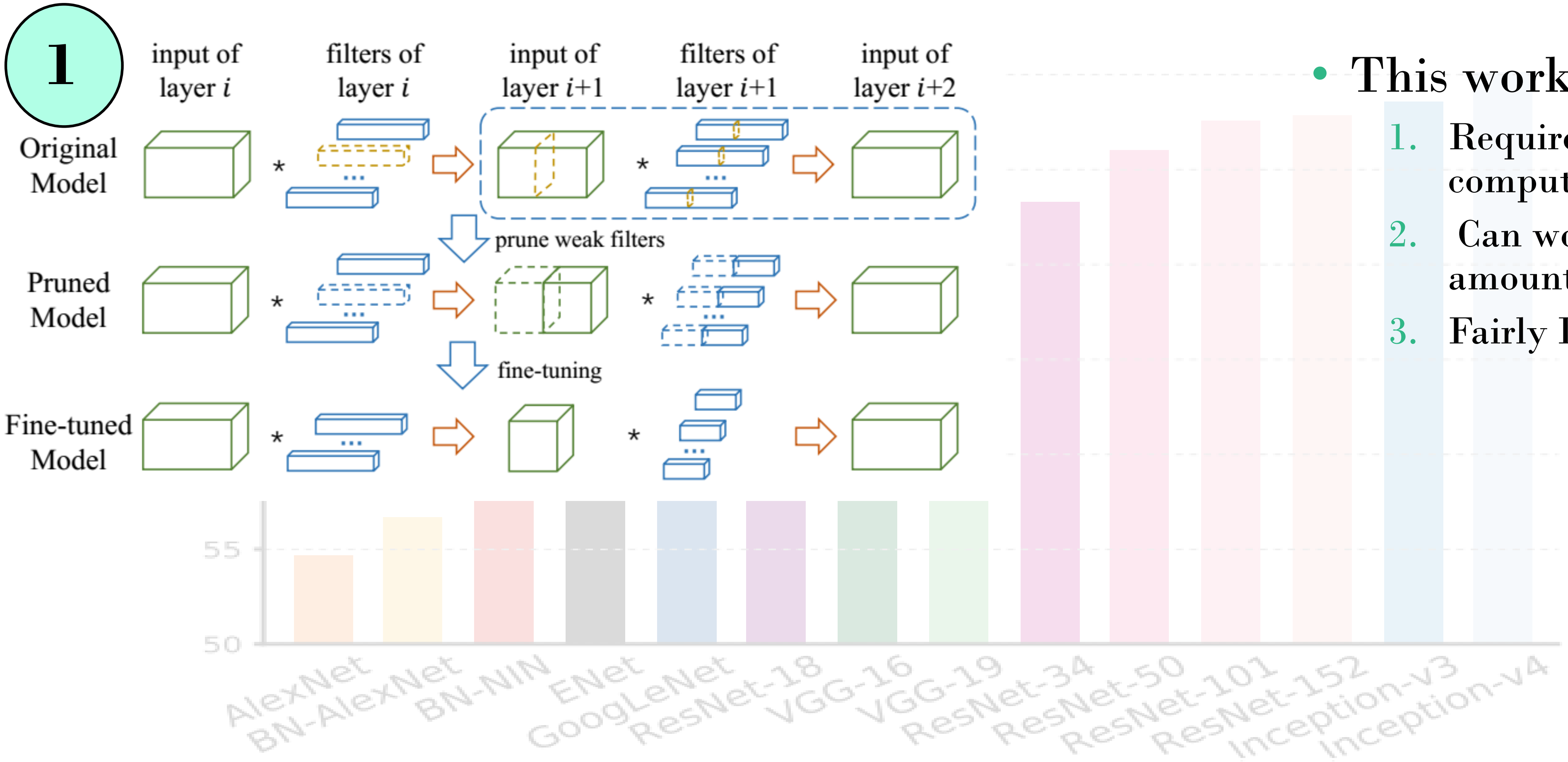
Deep Models have advanced image recognition field substantially



- This work
  1. Requires significantly less computational resources.
  2. Can work with available amount of data.
  3. Fairly Interpretable

Challenges

Deep Models have advanced image recognition field substantially

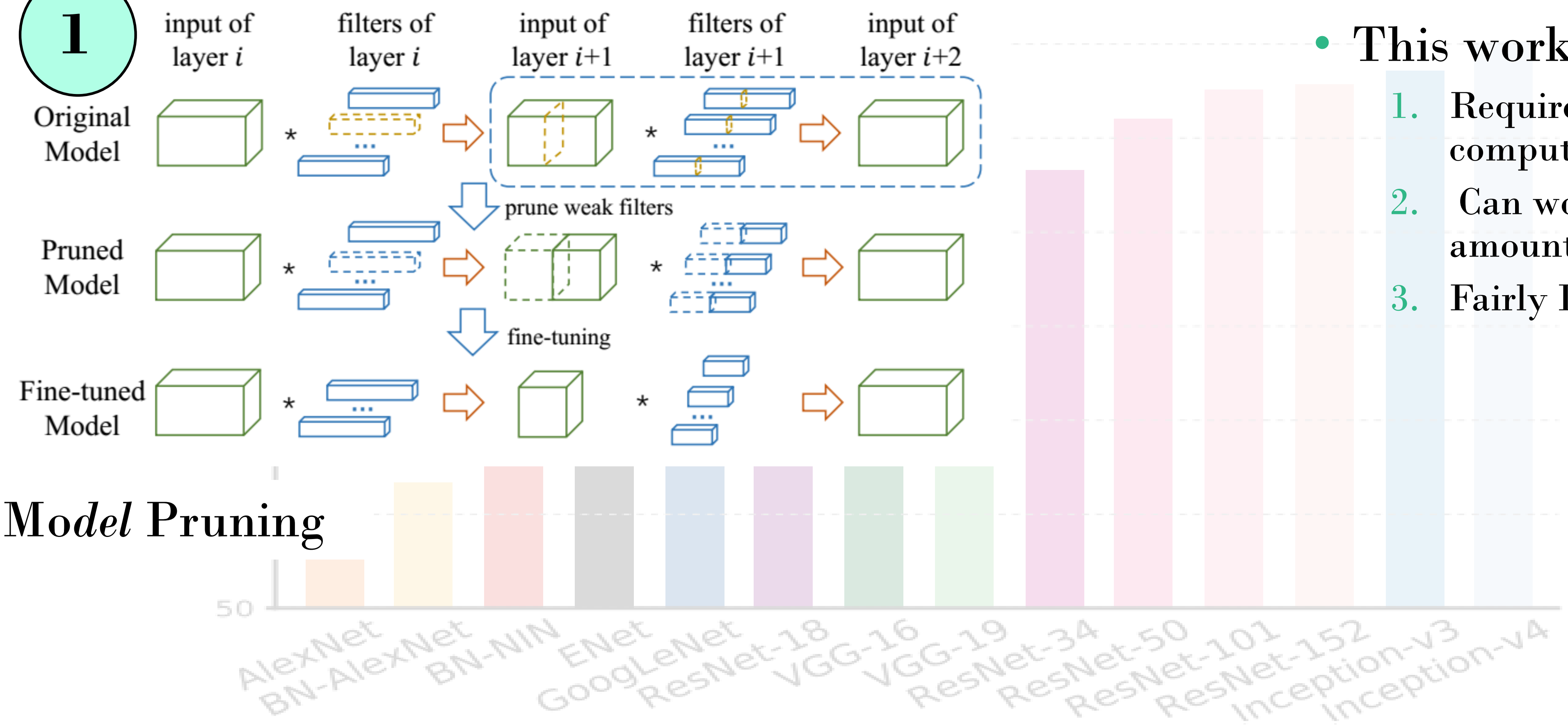


- This work
  1. Requires significantly less computational resources.
  2. Can work with available amount of data.
  3. Fairly Interpretable

Challenges

Deep Models have advanced image recognition field substantially

1



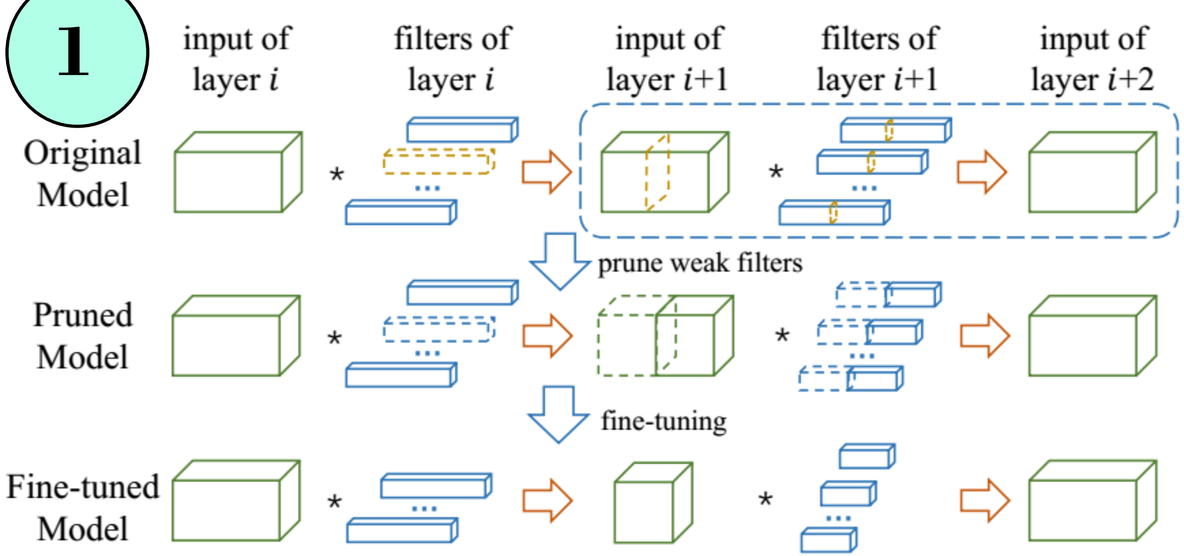
Model Pruning

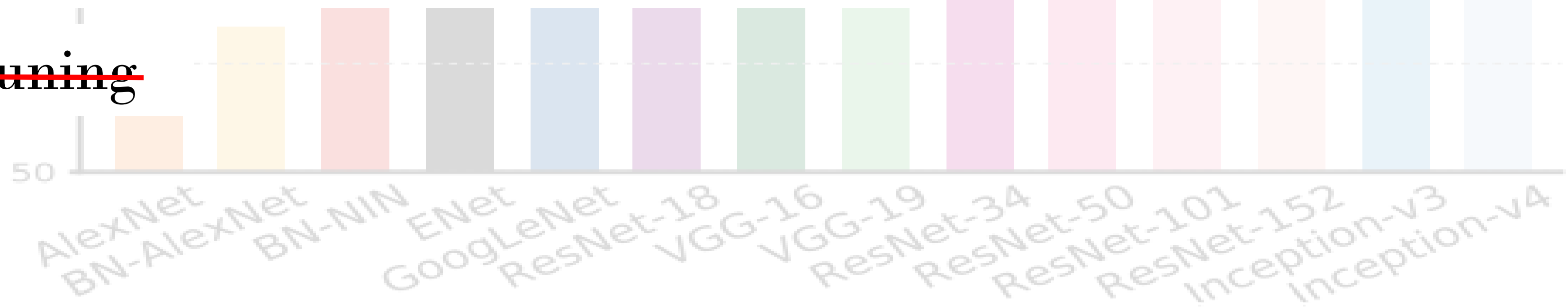
- This work
  1. Requires significantly less computational resources.
  2. Can work with available amount of data.
  3. Fairly Interpretable

Challenges

Deep Models have advanced image recognition field substantially

1





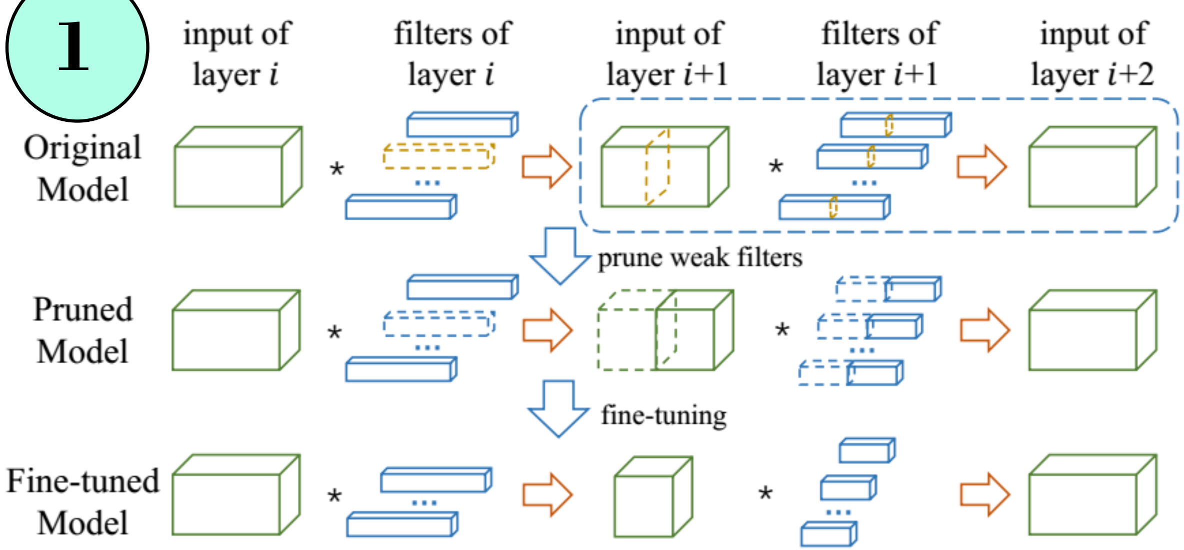
~~Model Pruning~~

- This work
  1. Requires significantly less computational resources.
  2. Can work with available amount of data.
  3. Fairly Interpretable

Challenges

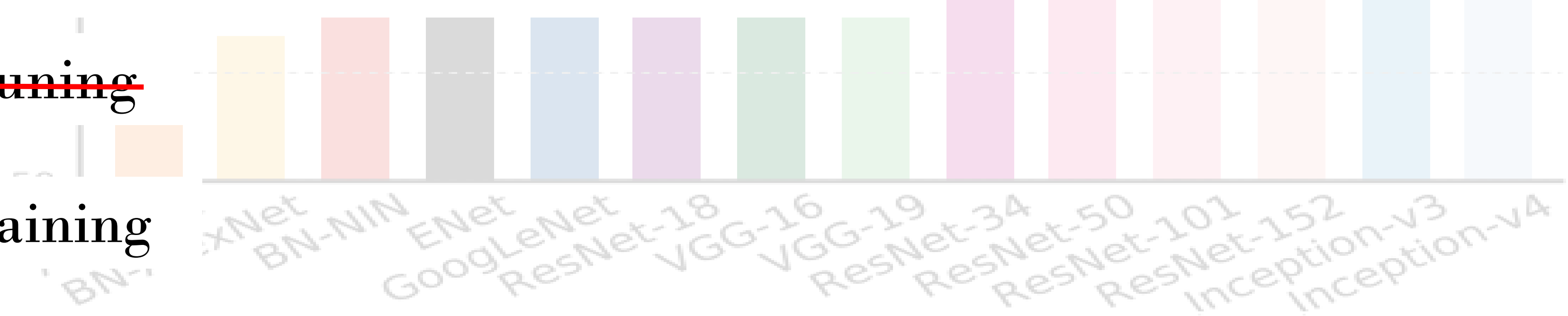
Deep Models have advanced image recognition field substantially

1



~~Model Pruning~~

Model Training

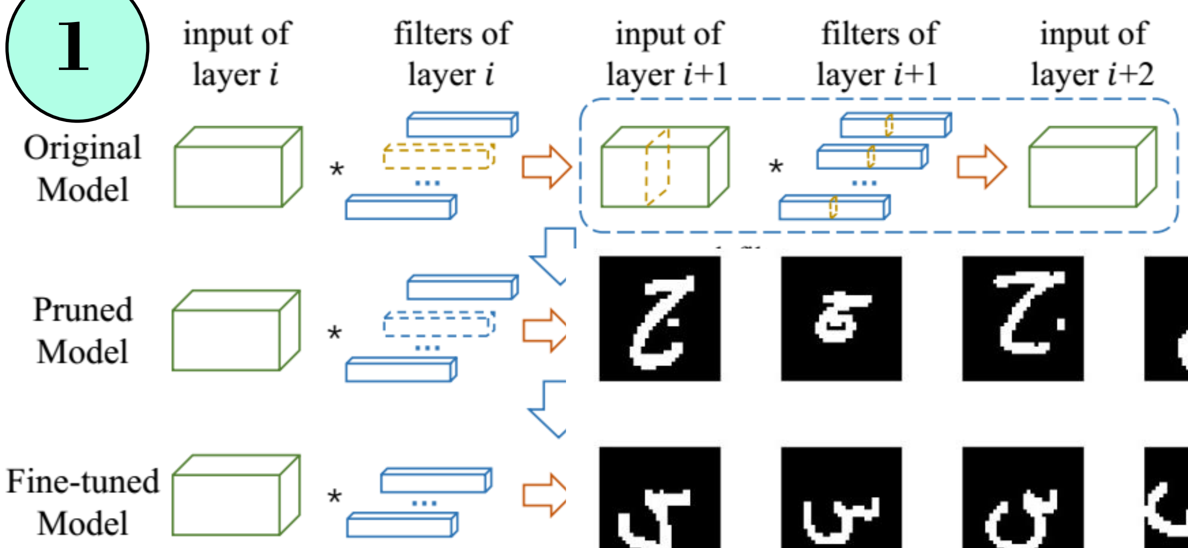


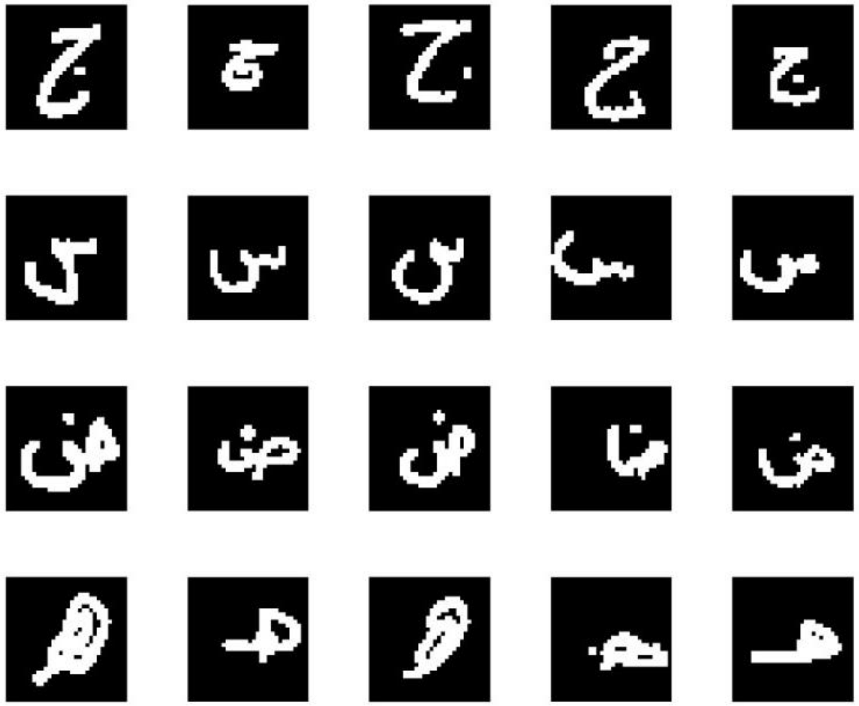
- This work
  1. Requires significantly less computational resources.
  2. Can work with available amount of data.
  3. Fairly Interpretable

Challenges

Deep Models have advanced image recognition field substantially

1





~~Model Pruning~~

Model Training

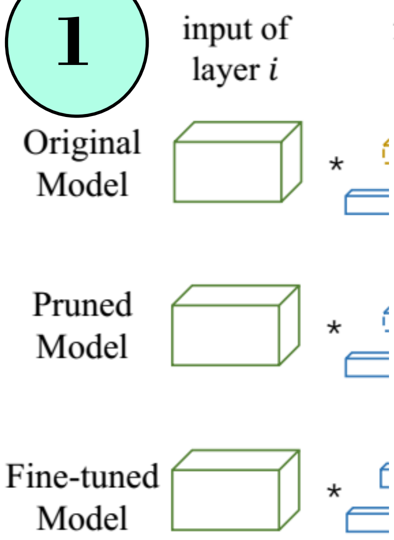
2

- This work
  1. Requires significantly less computational resources.
  2. Can work with available amount of data.
  3. Fairly Interpretable

Challenges

Deep Models have advanced image recognition field substantially

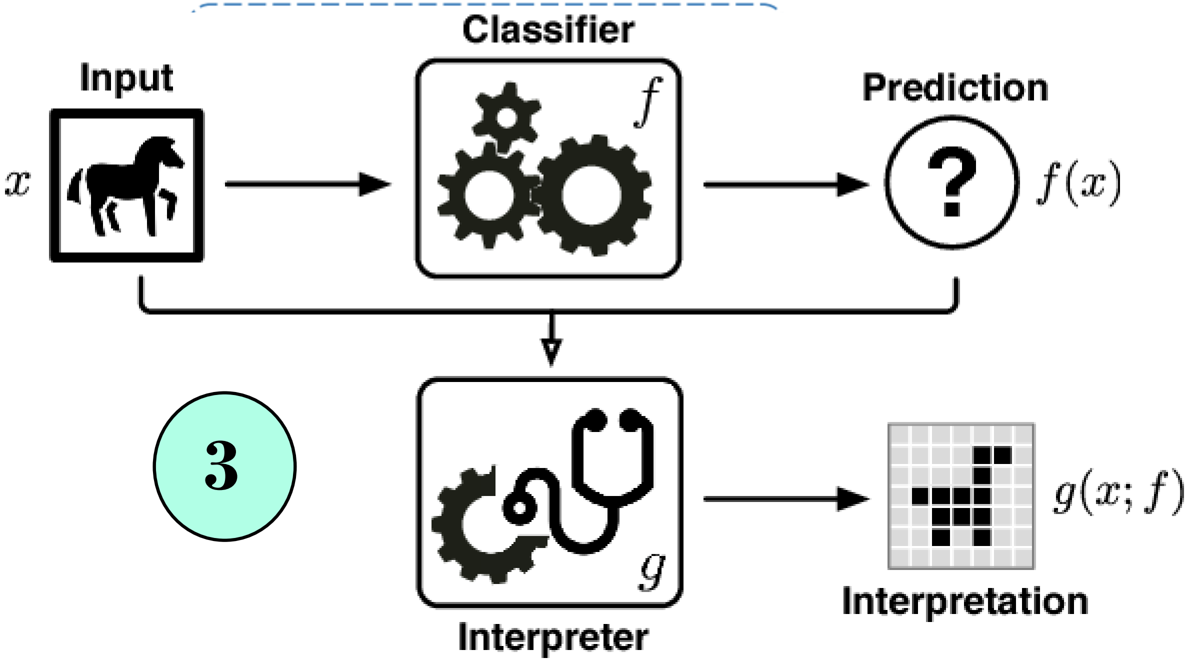
1



~~Model Pruning~~

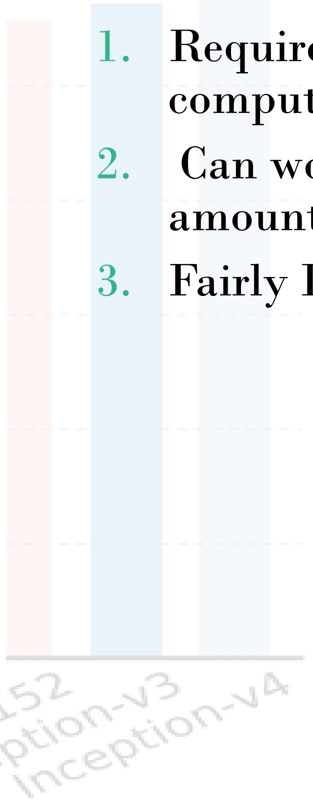
Model Training

3



2

- This work
  1. Requires significantly less computational resources.
  2. Can work with available amount of data.
  3. Fairly Interpretable

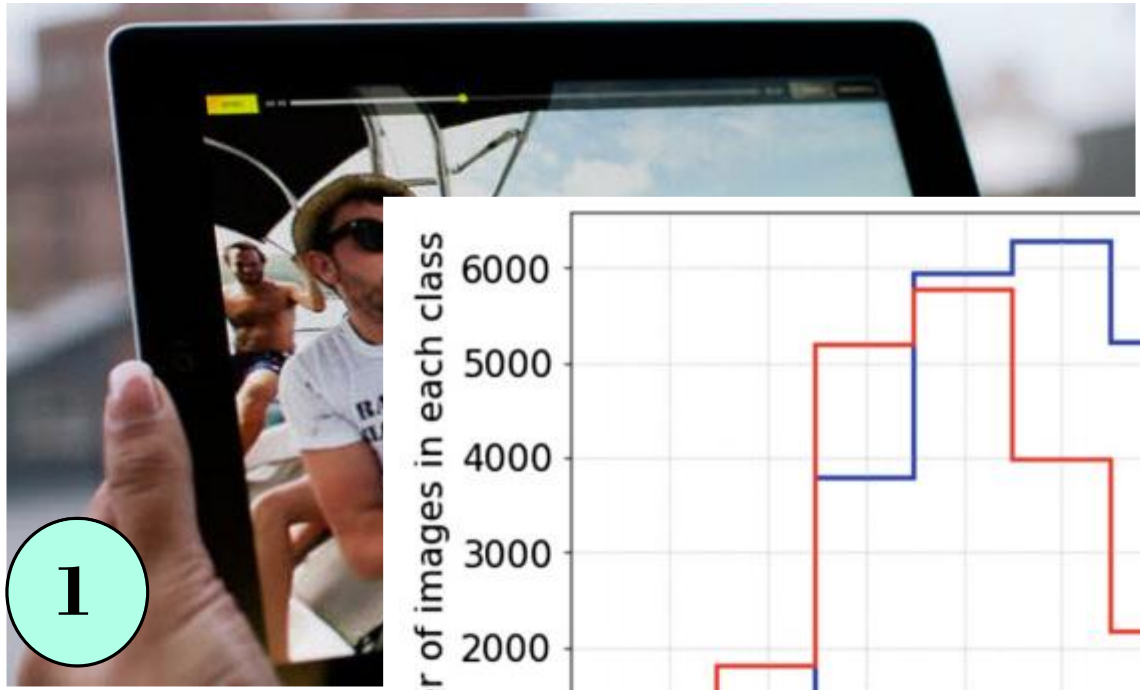


Importance

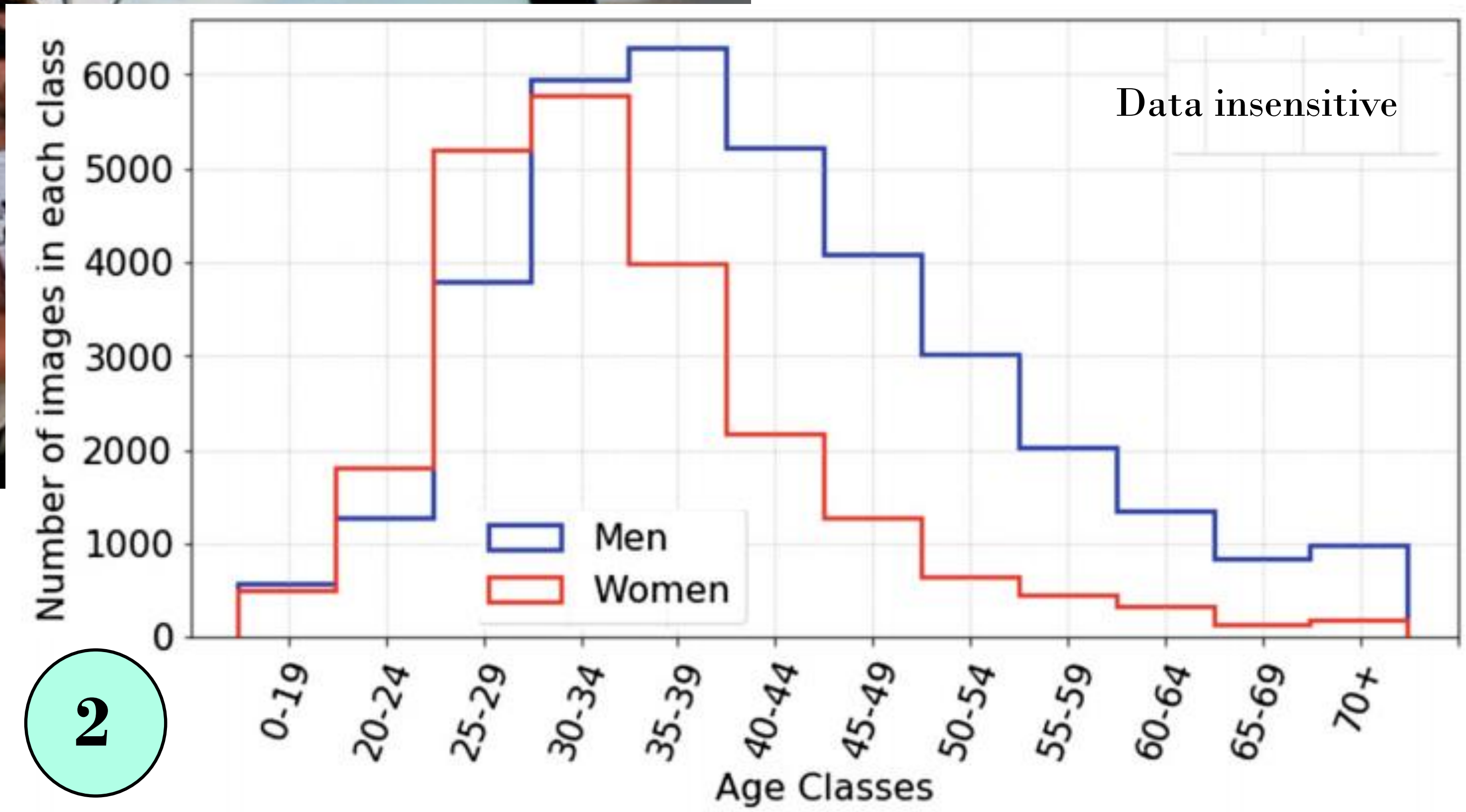
Importance

Environment Independent

Importance

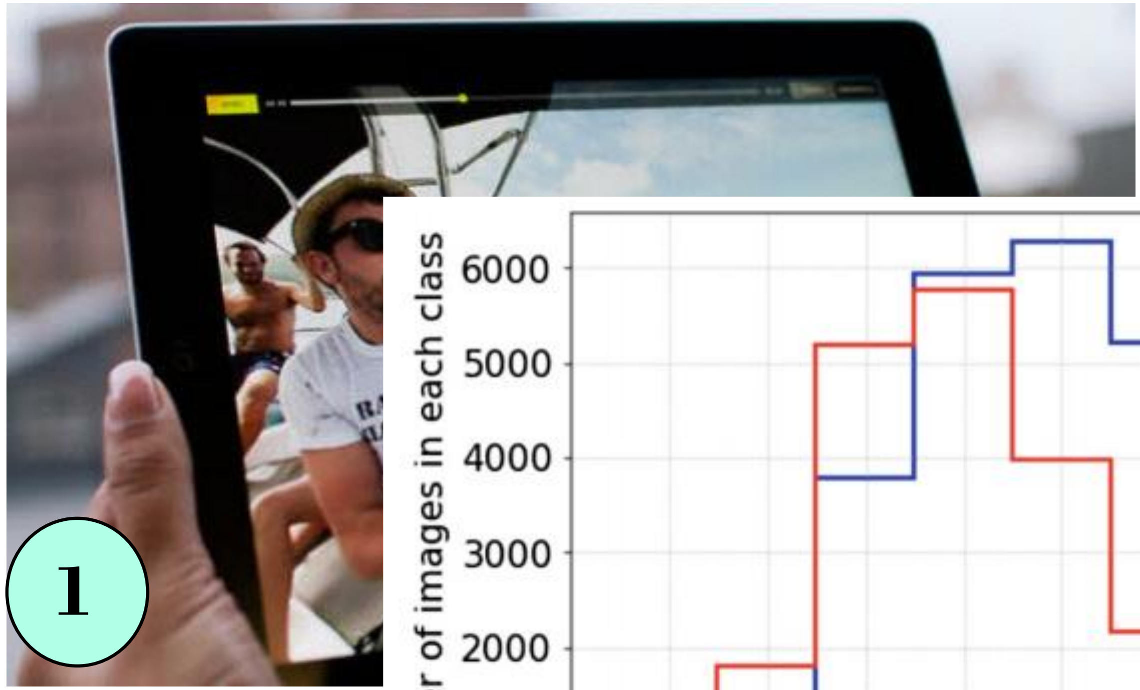


1

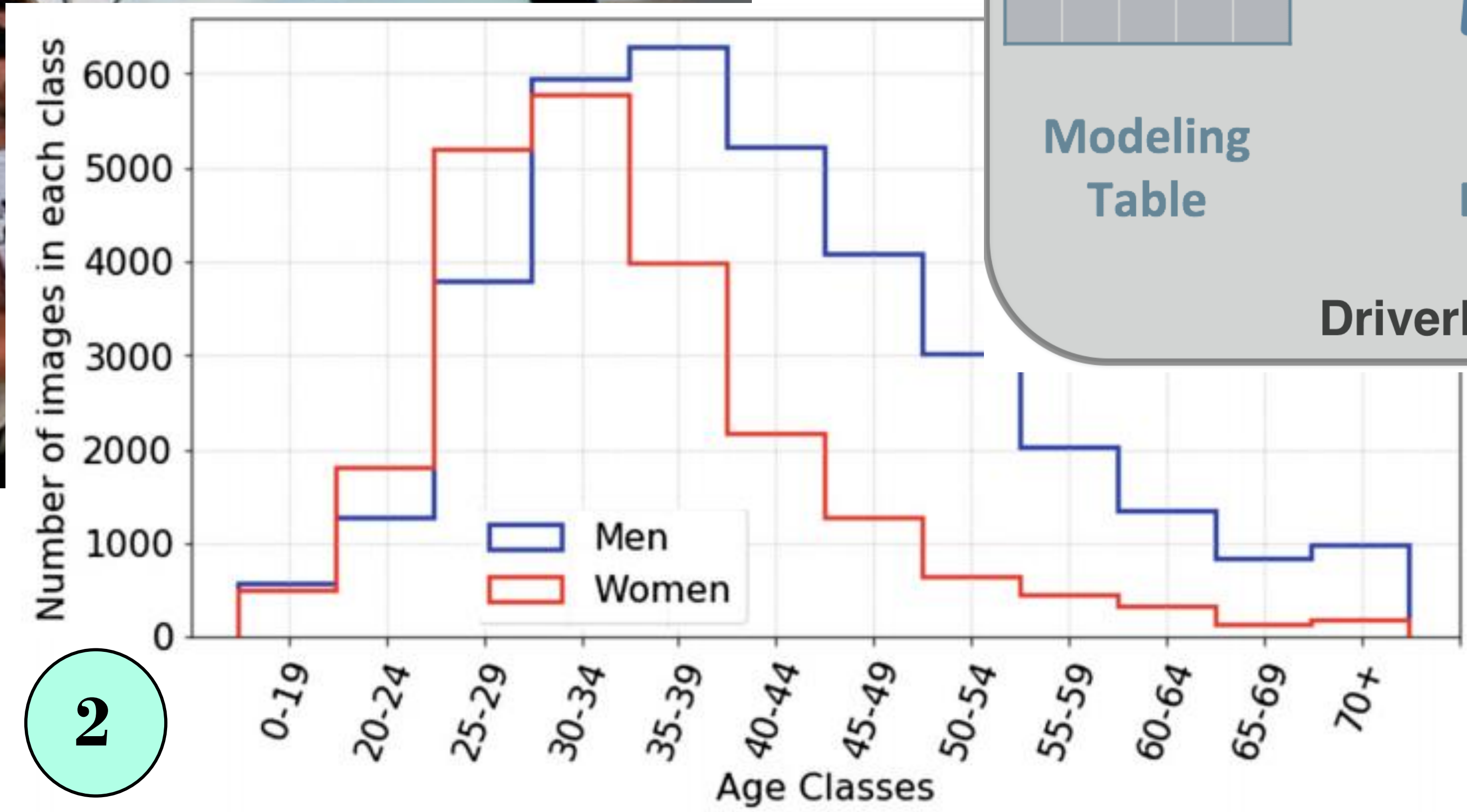


2

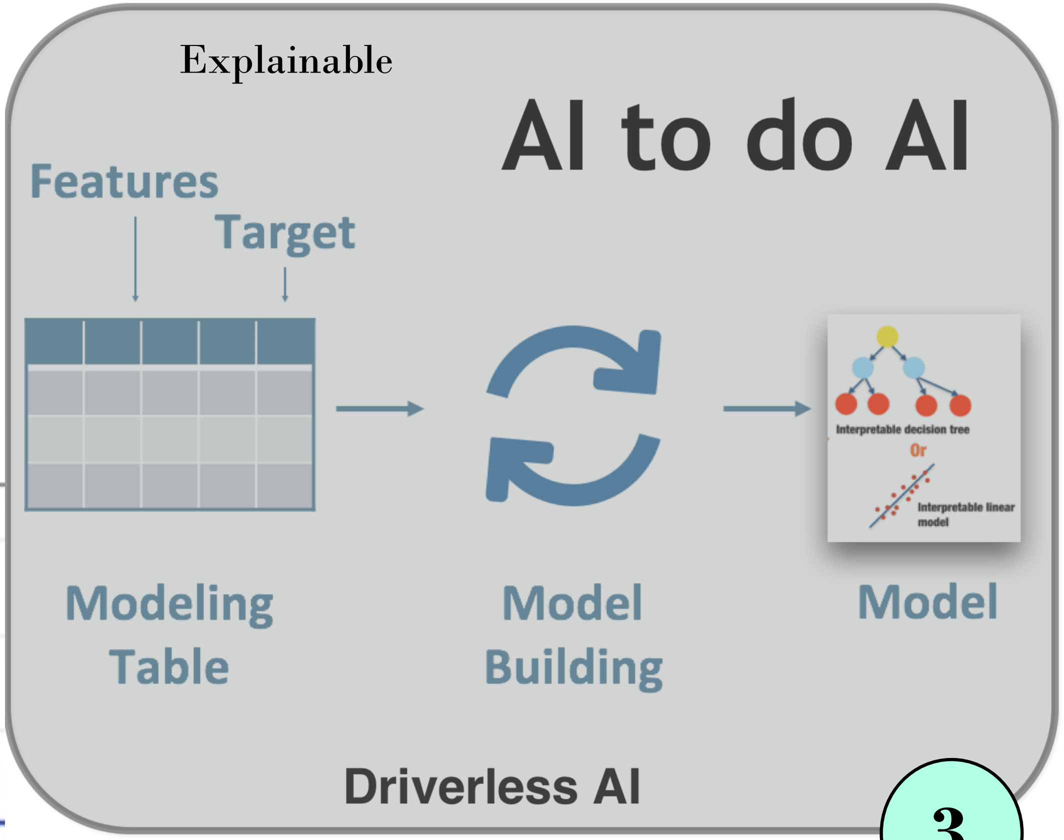
Importance



1



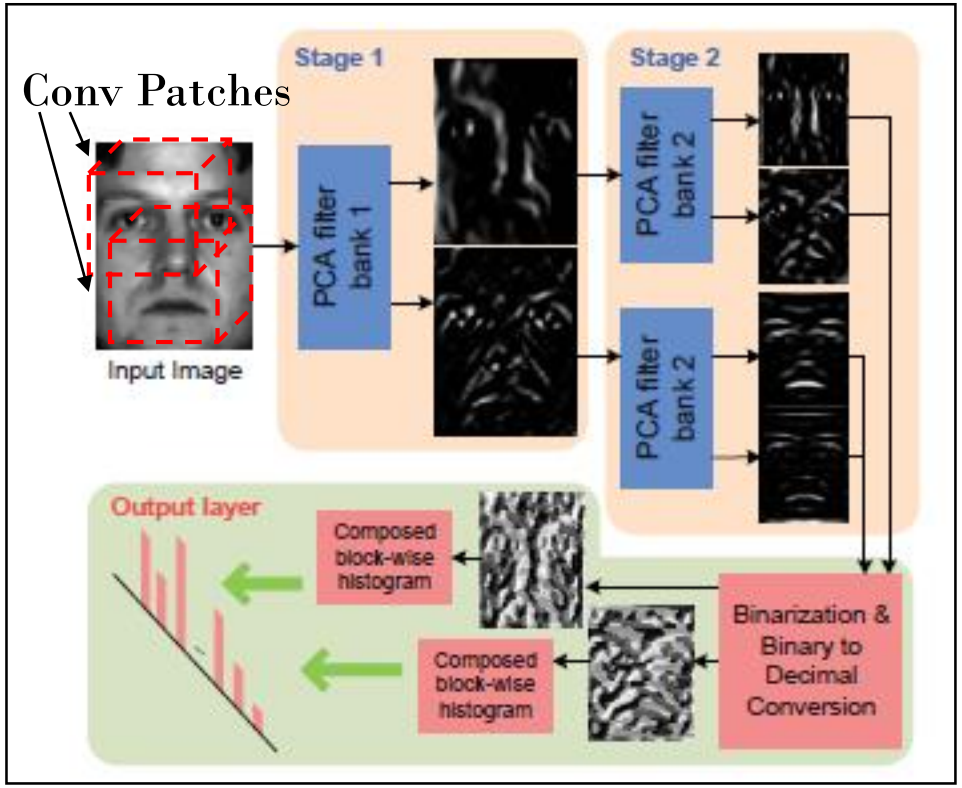
2



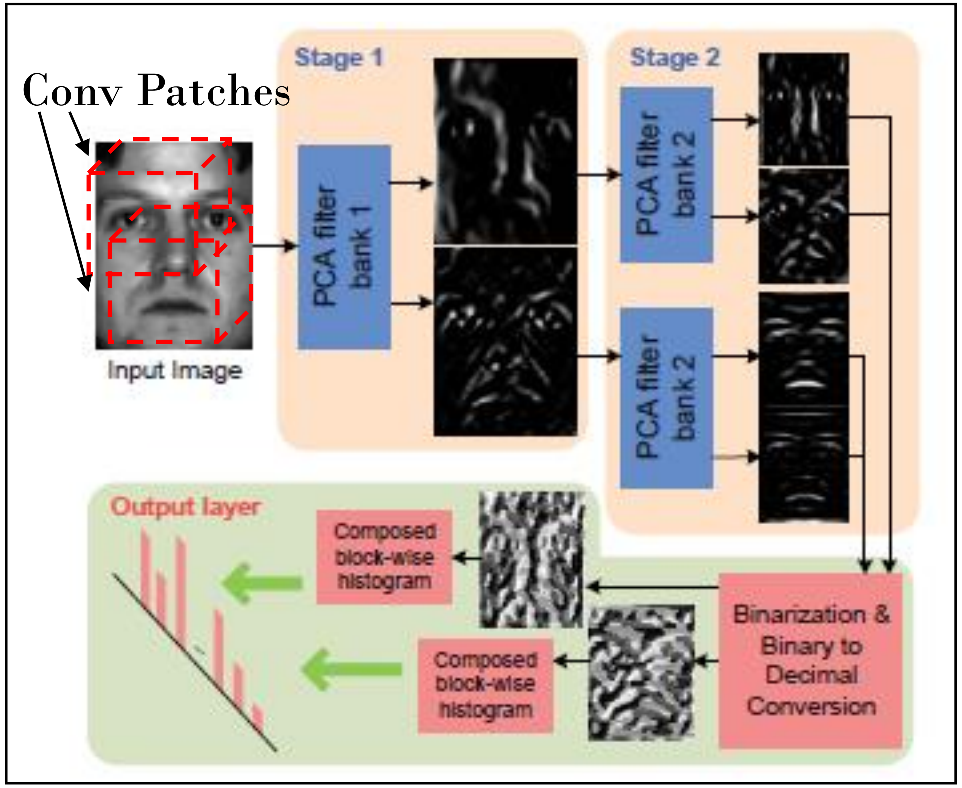
3

Existing Solution

PCANet



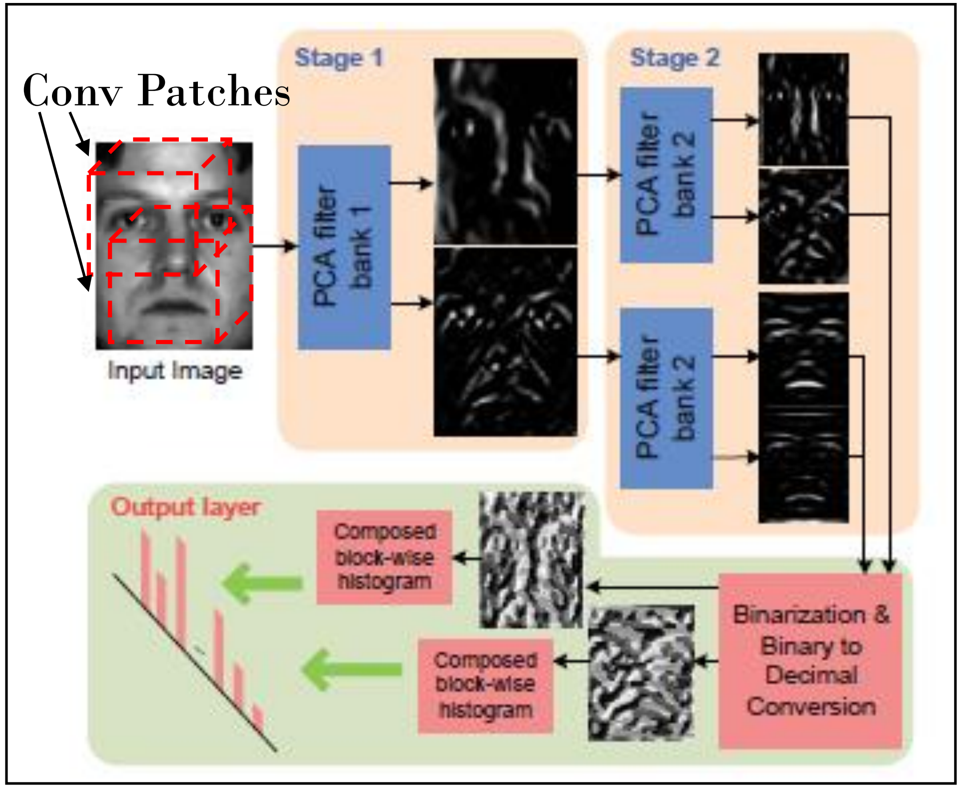
Existing Solution
PCANet



- Advantages
  1. Light Weight Deep Network.
  2. Acceptable performance with less data.
  3. Unsupervised feature extractor
     - Less sensitive to data perturbation

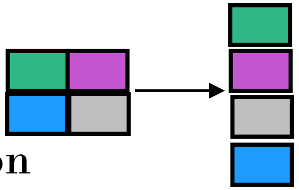
Existing Solution

PCANet



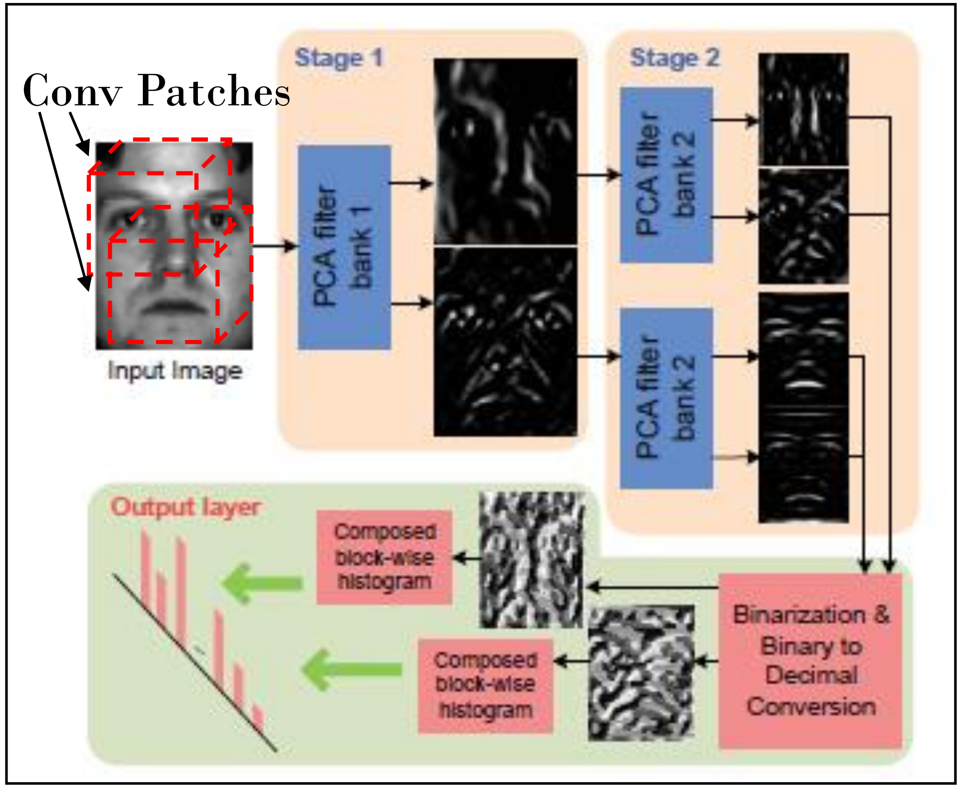
- Advantages
  1. Light Weight Deep Network.
  2. Acceptable performance with less data.
  3. Unsupervised feature extractor
     - Less sensitive to data perturbation

- Drawbacks
  1. Vectorizes each patch
     - Destroys spatial information



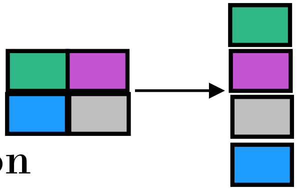
Existing Solution

PCANet

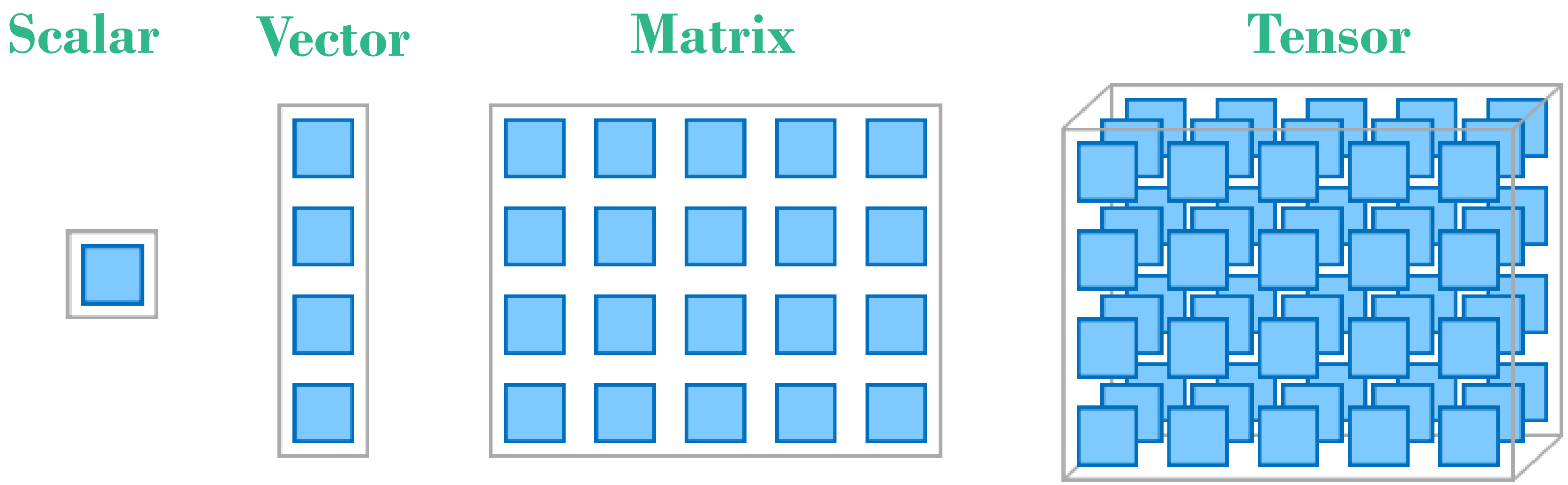


- Advantages
  1. Light Weight Deep Network.
  2. Acceptable performance with less data.
  3. Unsupervised feature extractor
     - Less sensitive to data perturbation

- Drawbacks
  1. Vectorizes each patch
     - Destroys spatial information
  2. Patch matrix becomes tall/wide.
     - Requires more computational resources
     - Better algorithms

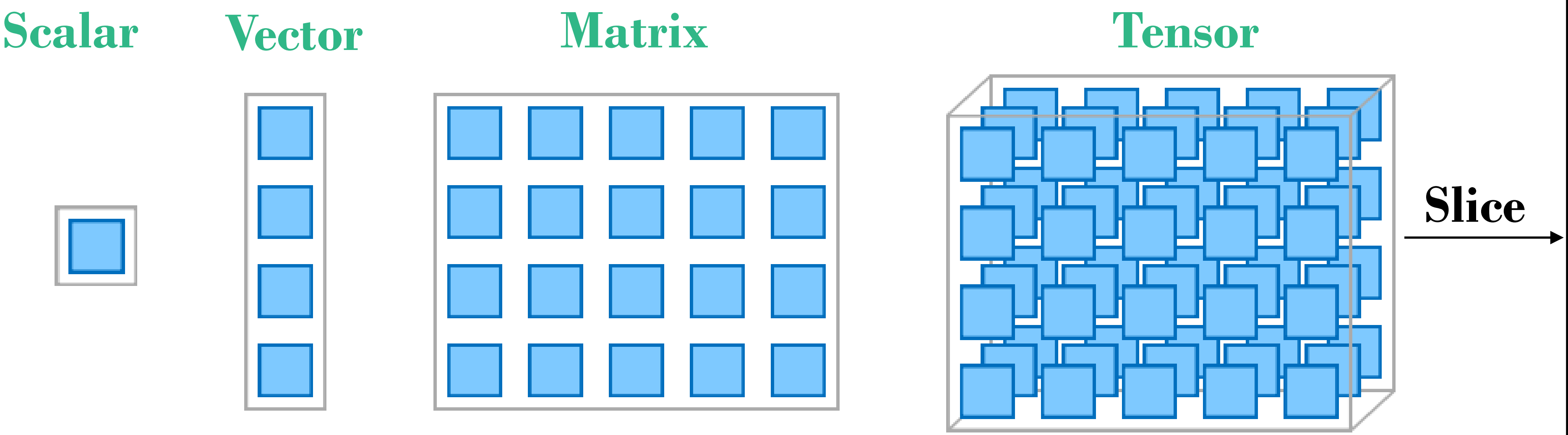


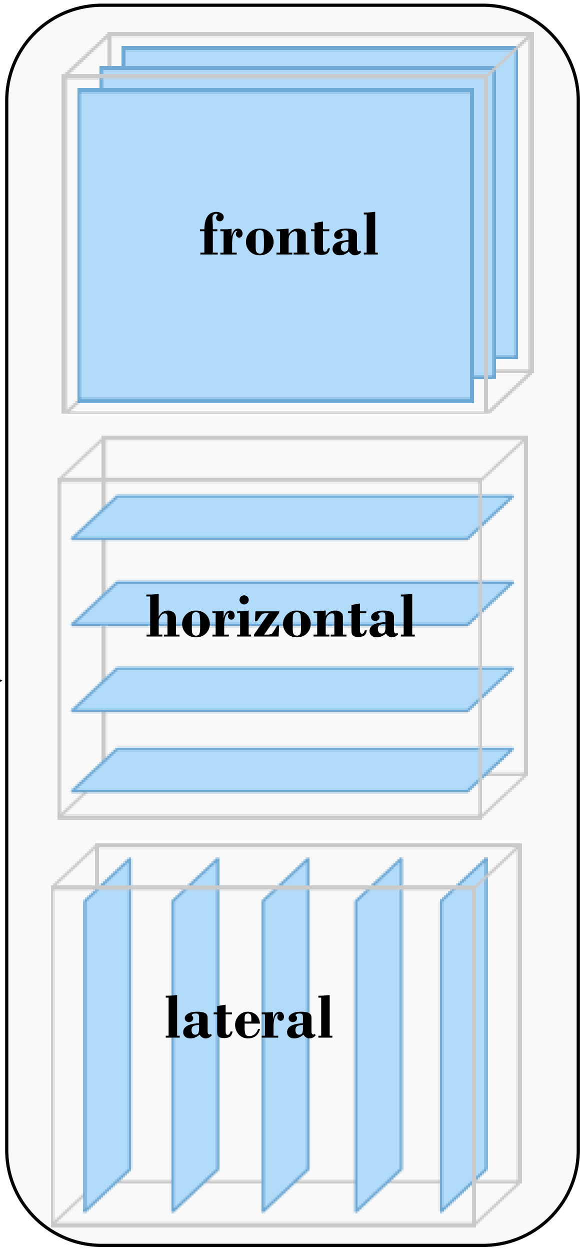
Tensor Network, Preliminaries



$$x \in \mathbb{R}, \boldsymbol{x} \in \mathbb{R}^4, \boldsymbol{X} \in \mathbb{R}^{4\times5}, \boldsymbol{\mathcal{X}} \in \mathbb{R}^{4\times5\times3}$$

Tensor Network, Preliminaries



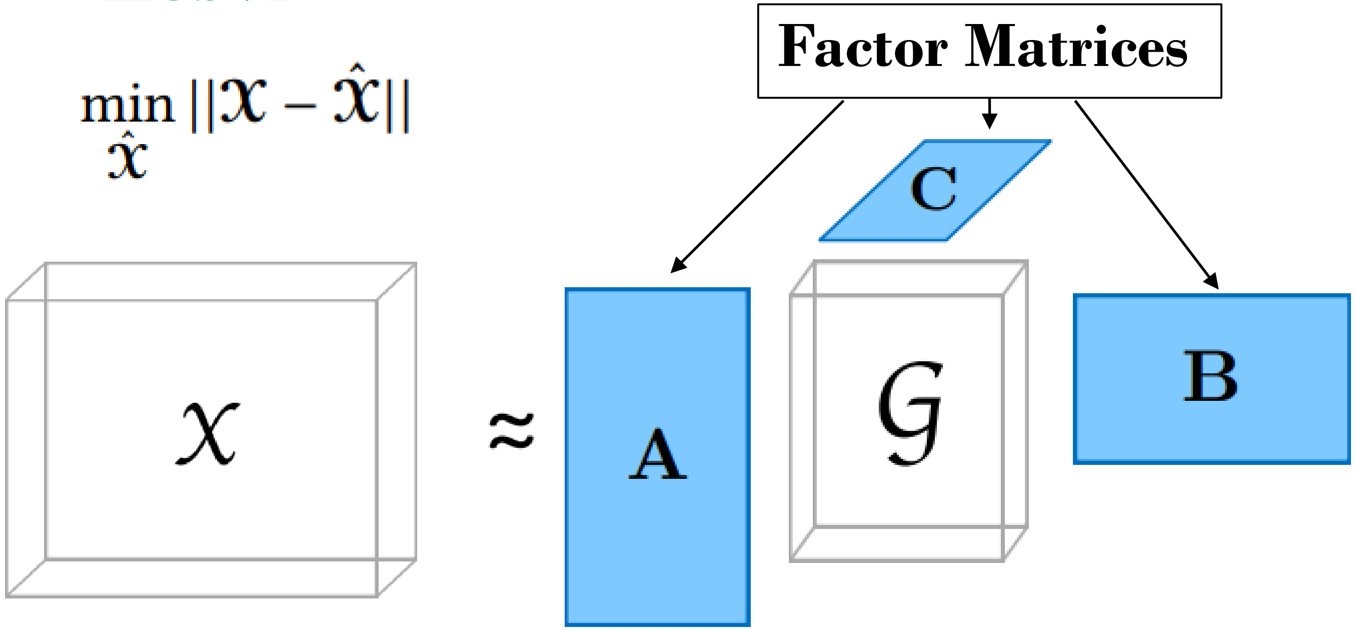


$$x \in \mathbb{R}, \boldsymbol{x} \in \mathbb{R}^4, \boldsymbol{X} \in \mathbb{R}^{4\times5}, \boldsymbol{\mathcal{X}} \in \mathbb{R}^{4\times5\times3}$$

TDNet

HOSVD

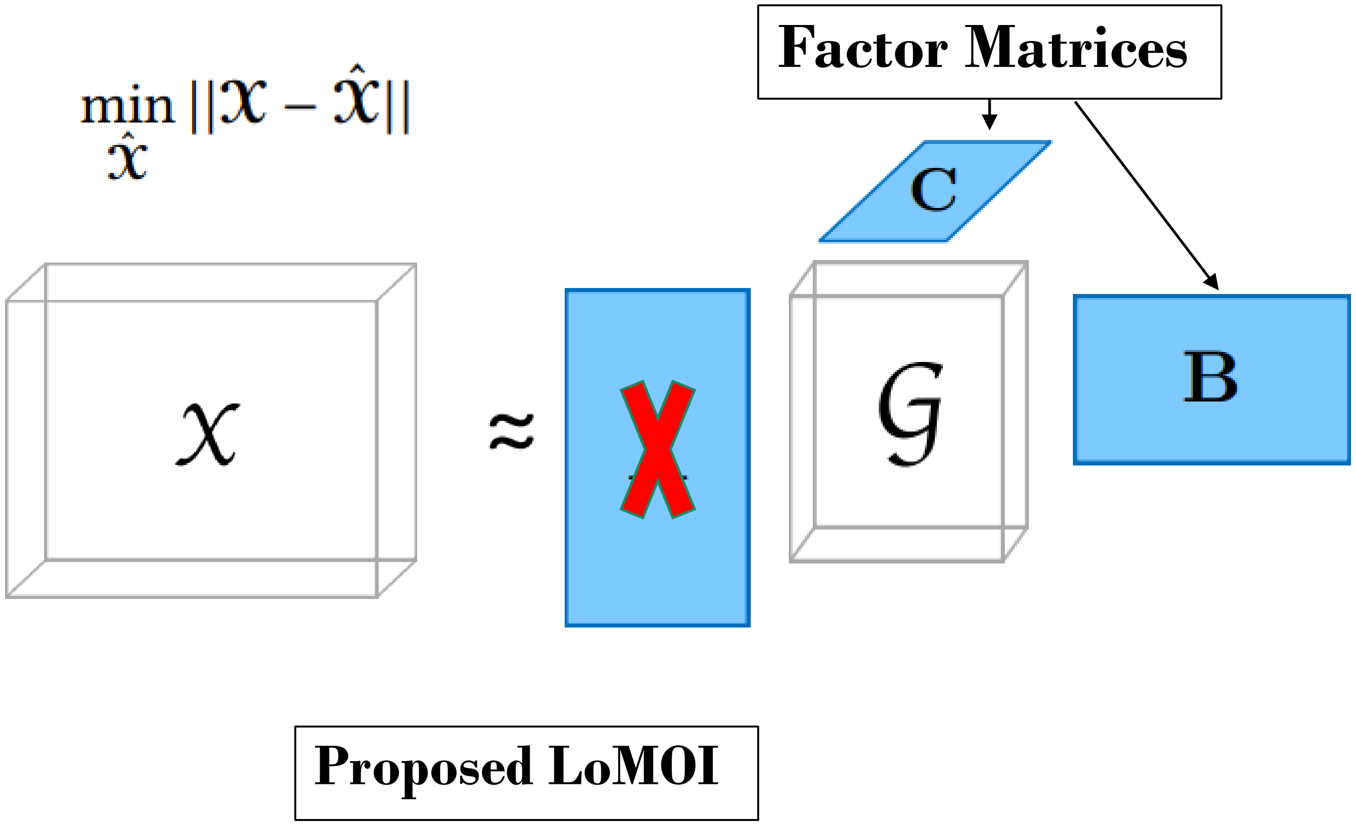
$$\min_{\hat{\mathcal{X}}} ||\mathcal{X} - \hat{\mathcal{X}}||$$



Tucker Decomposition

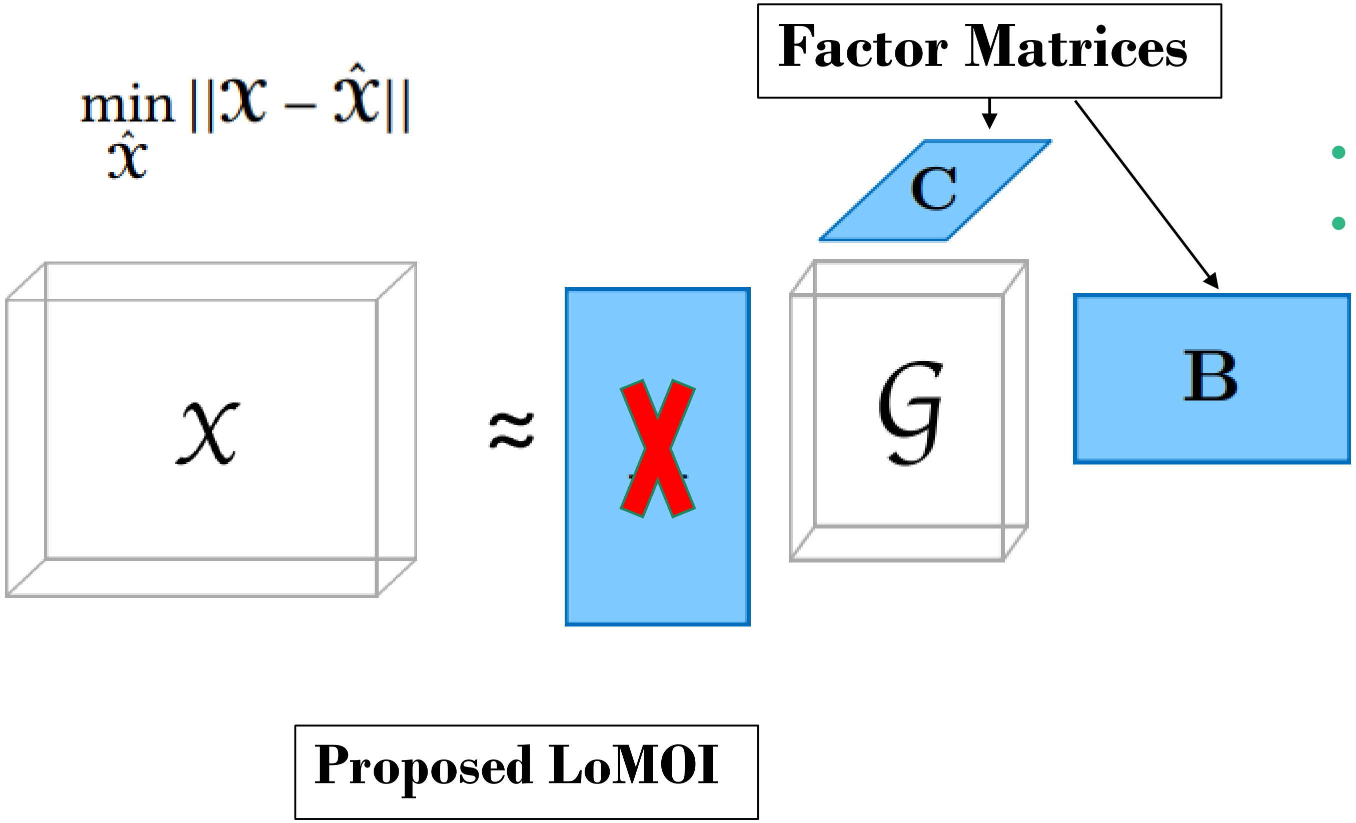
TDNet

$$\min_{\hat{\mathcal{X}}} ||\mathcal{X} - \hat{\mathcal{X}}||$$



TDNet

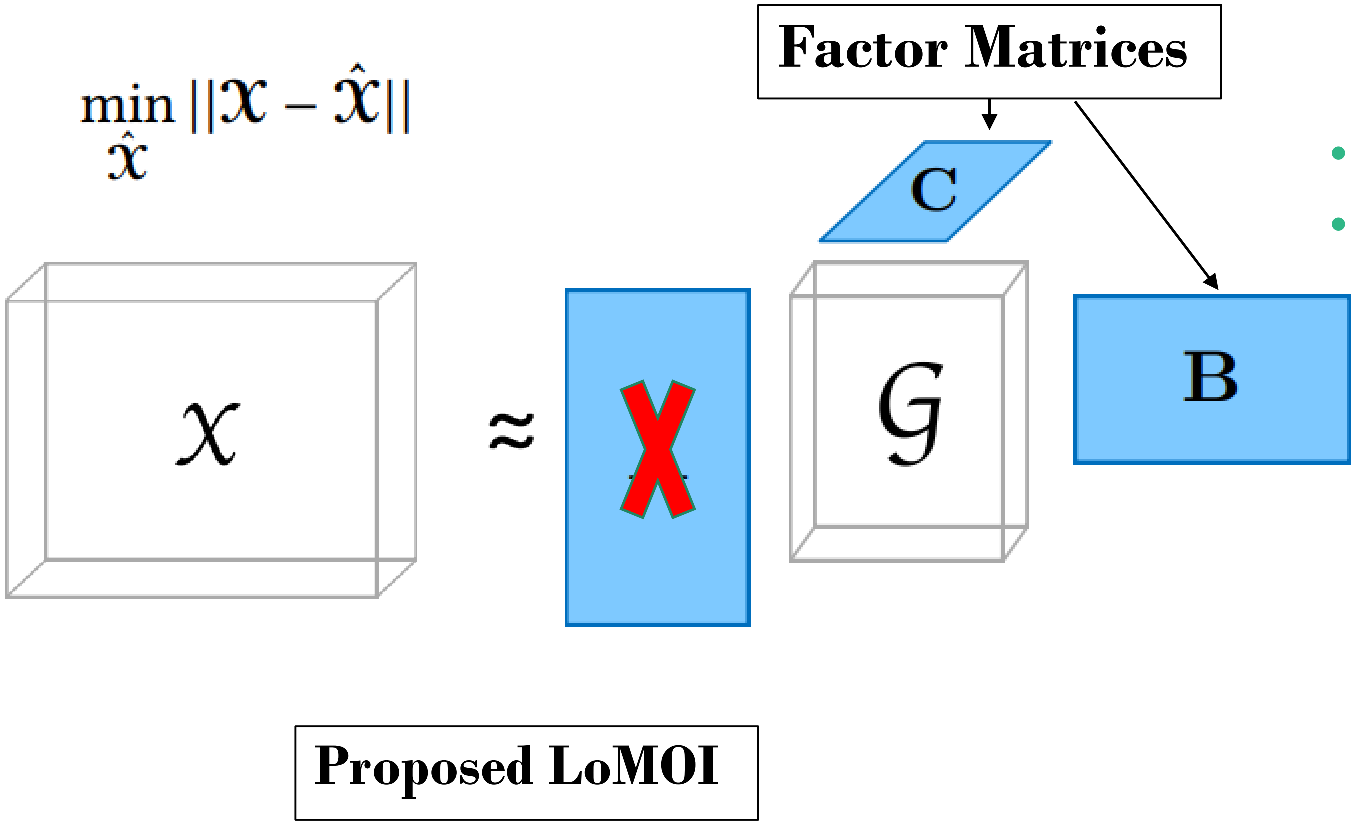
$$\min_{\hat{\mathcal{X}}} ||\mathcal{X} - \hat{\mathcal{X}}||$$



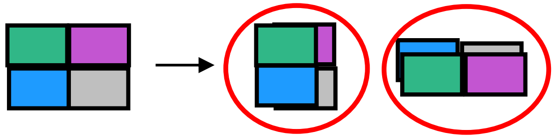
- Learns from minutiae view of the data.
- Preserves spatial structure in the data

TDNet

$$\min_{\hat{\mathcal{X}}} ||\mathcal{X} - \hat{\mathcal{X}}||$$

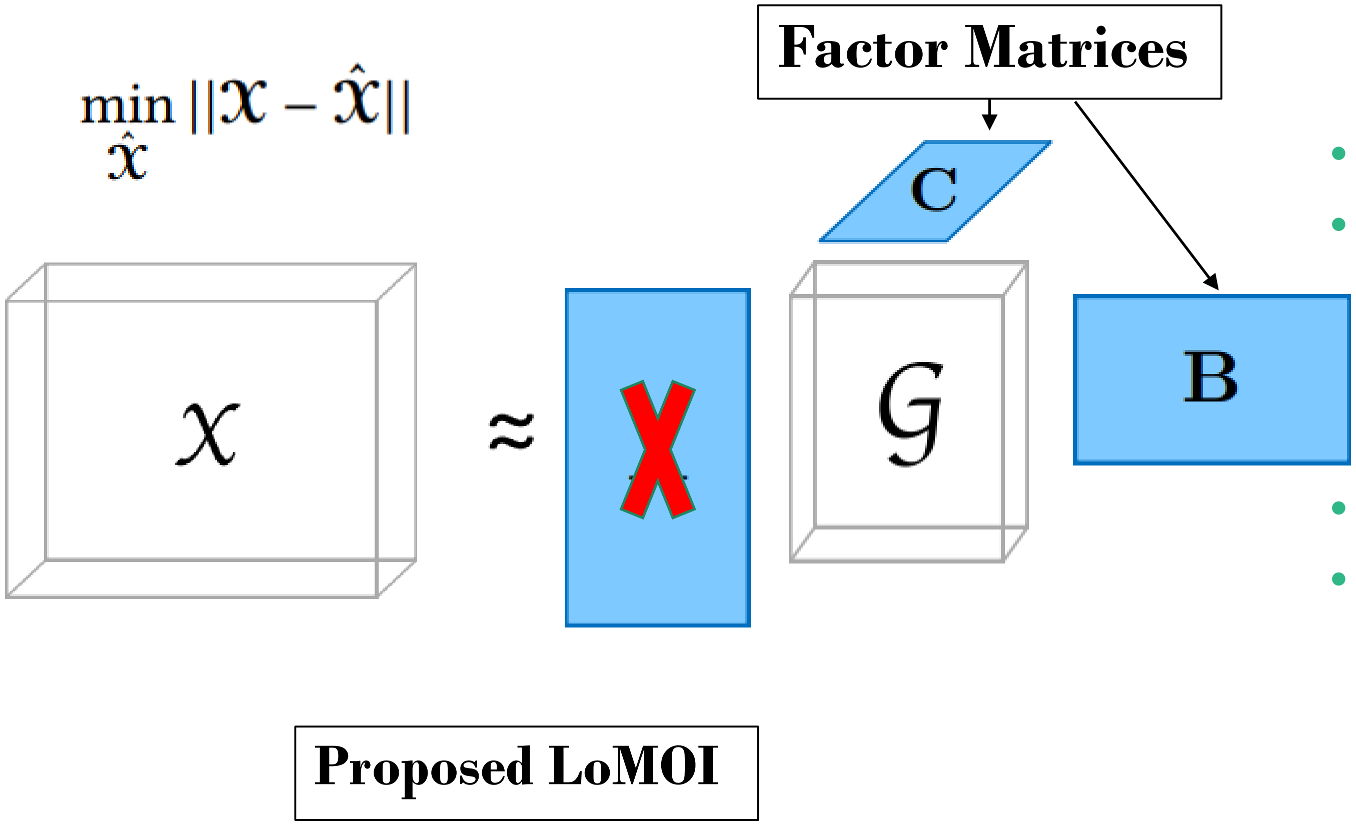


- Learns from minutiae view of the data.
- Preserves spatial structure in the data

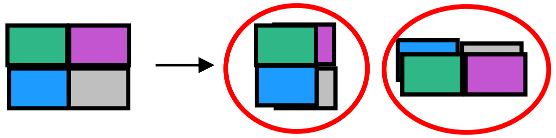


TDNet

$$\min_{\hat{\mathcal{X}}} ||\mathcal{X} - \hat{\mathcal{X}}||$$

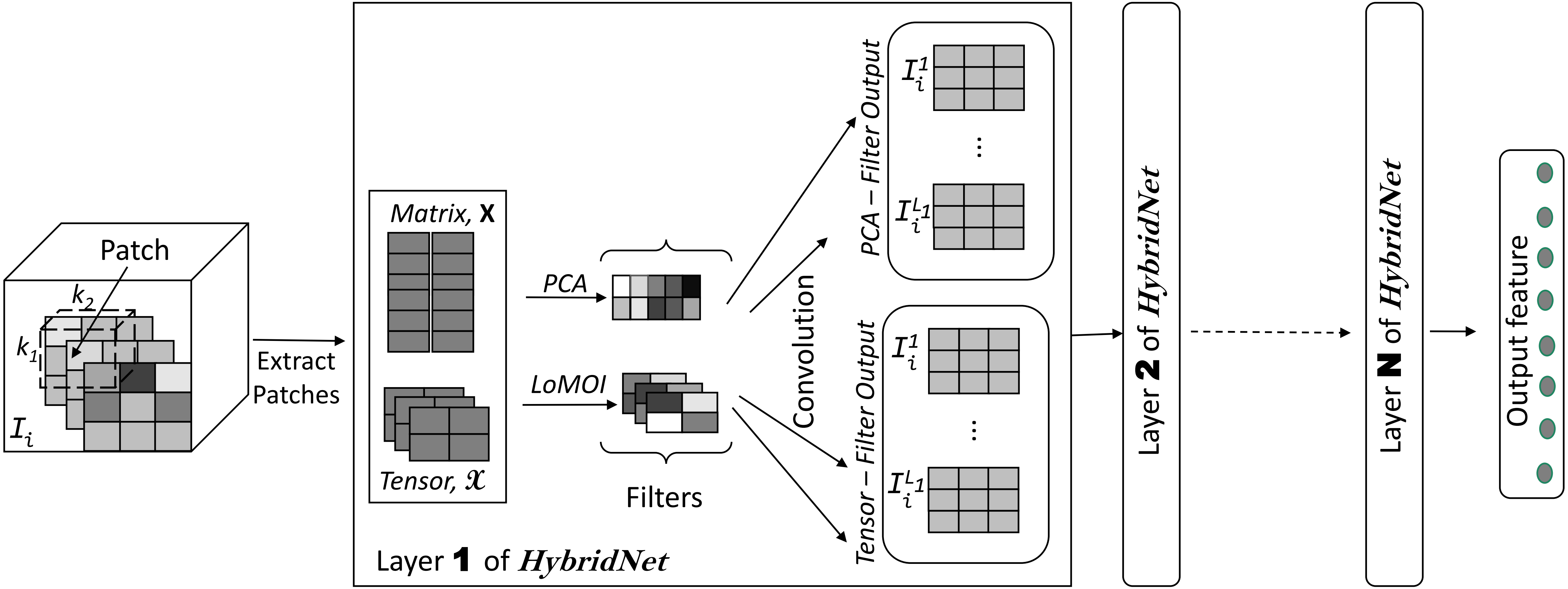


- Learns from minutiae view of the data.
- Preserves spatial structure in the data

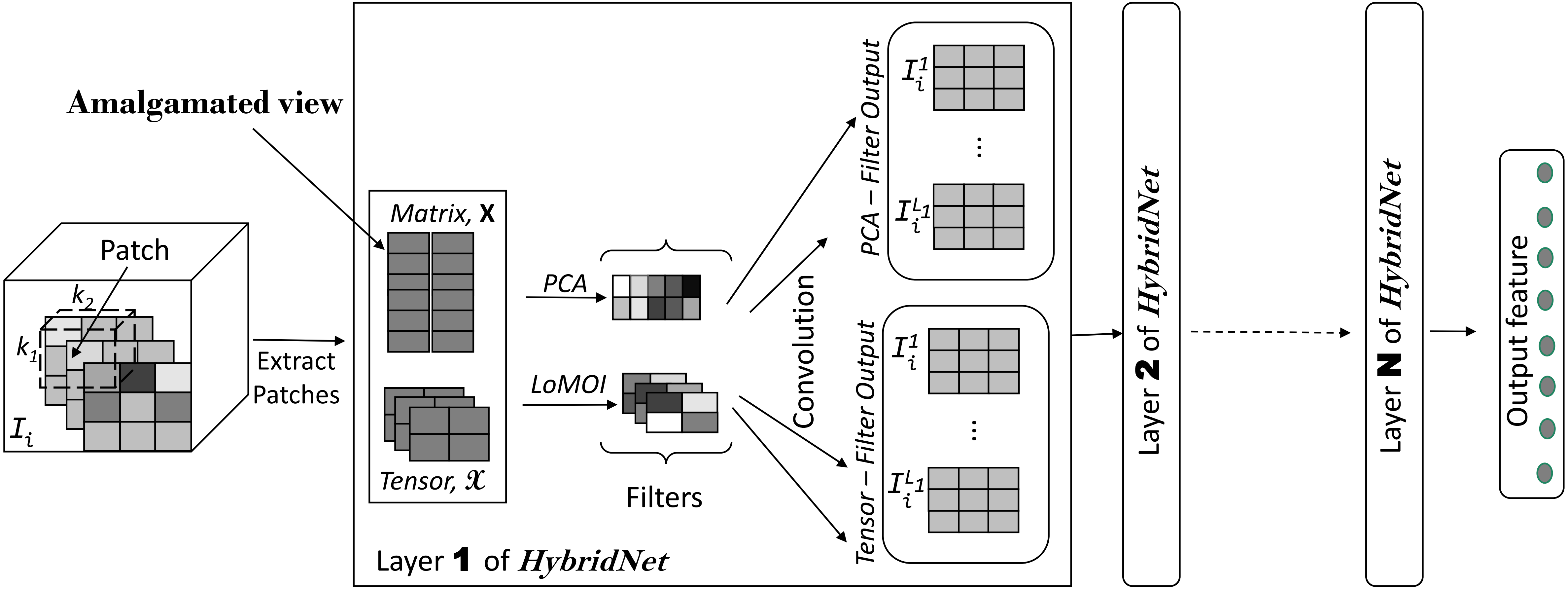


- Computationally less expensive than PCA.
- Captures variations in each mode independently.

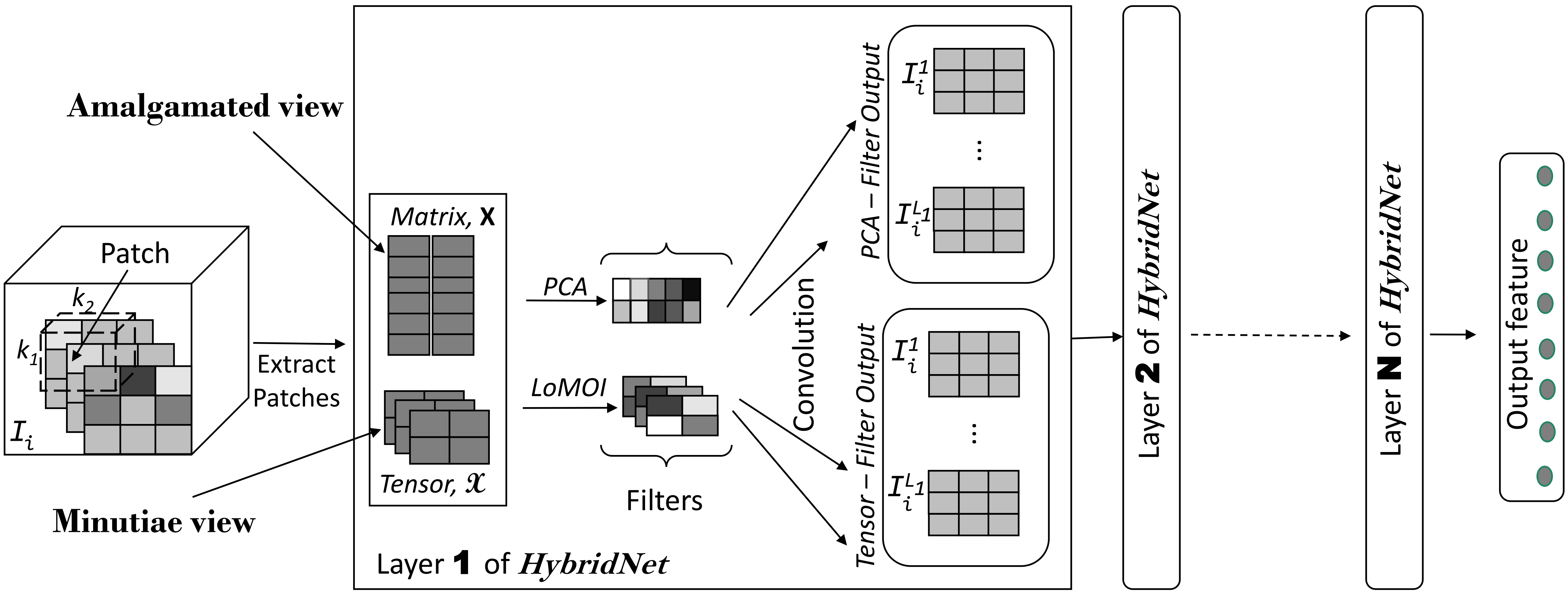
Hybrid Networks



Hybrid Networks

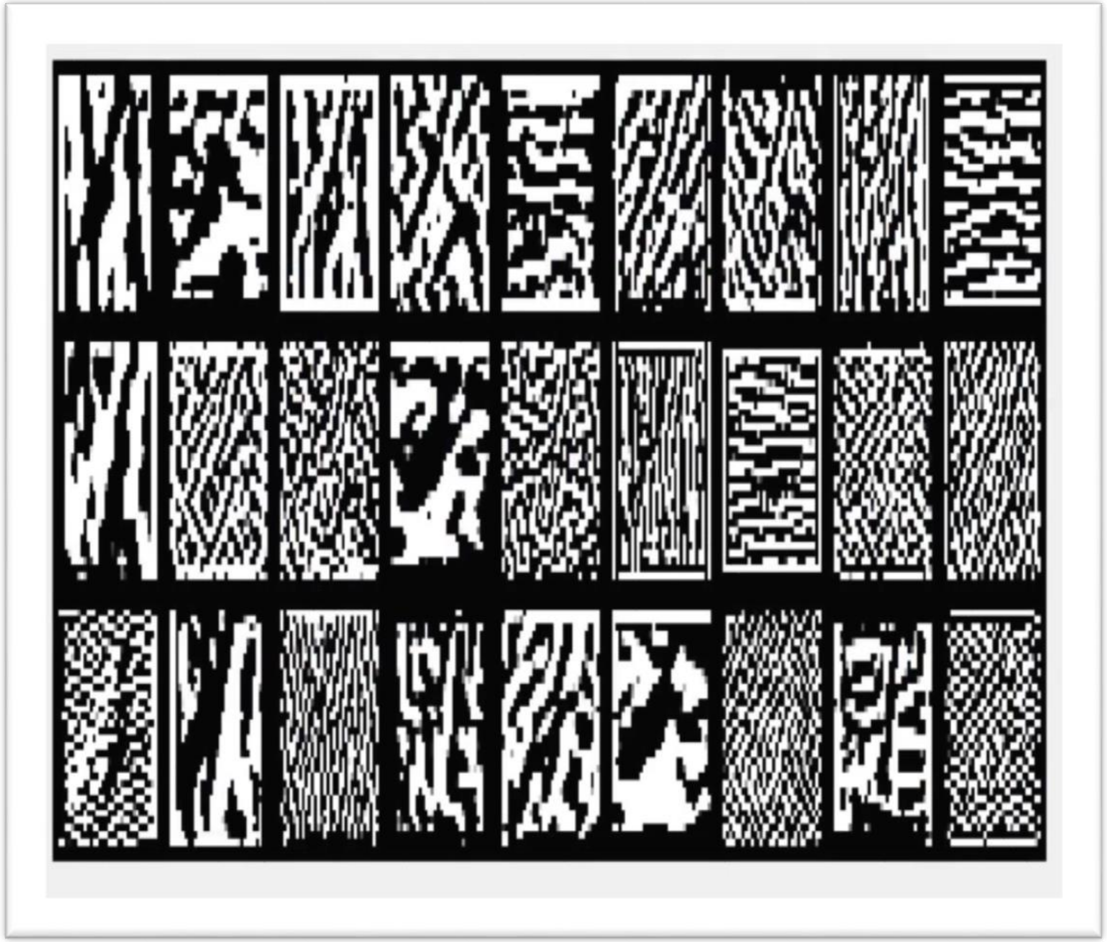


Hybrid Networks



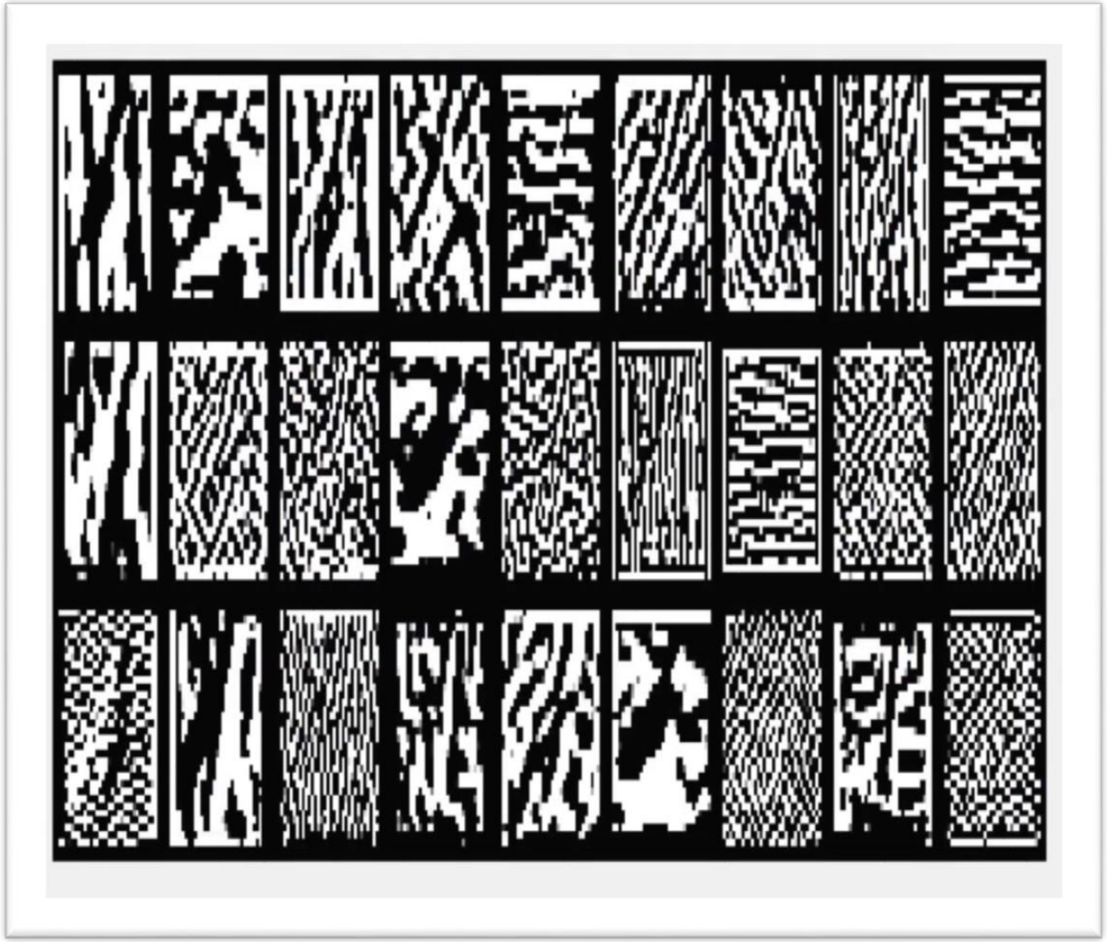
Output from the first layer in Hybrid Networks

PCANet

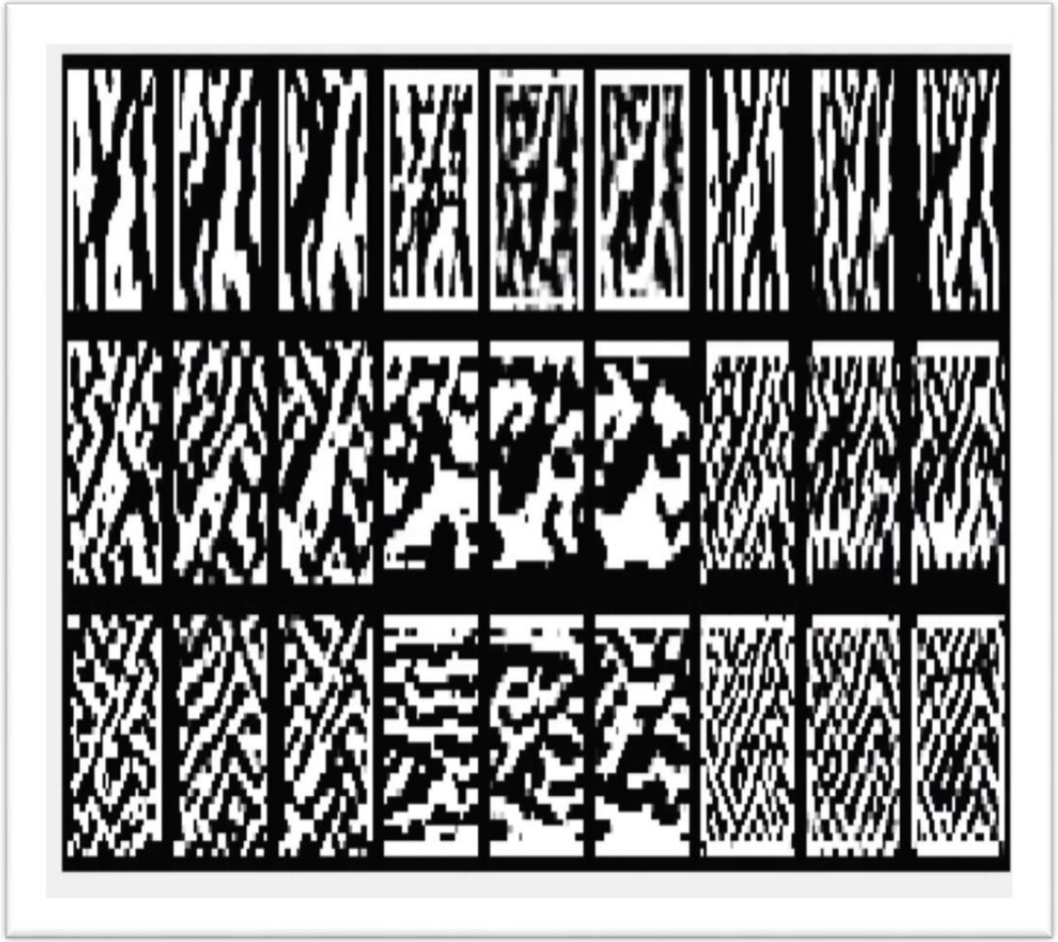


Output from the first layer in Hybrid Networks

PCANet

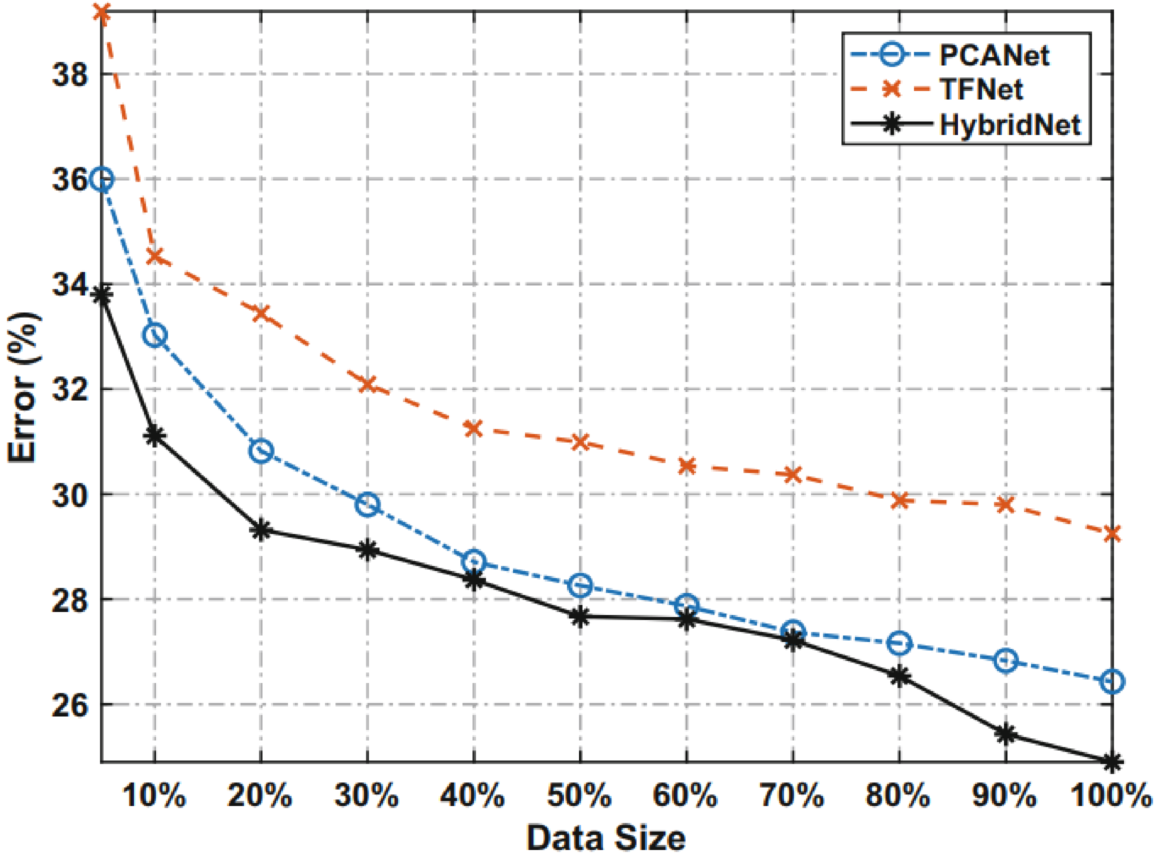


TensorNet



Results

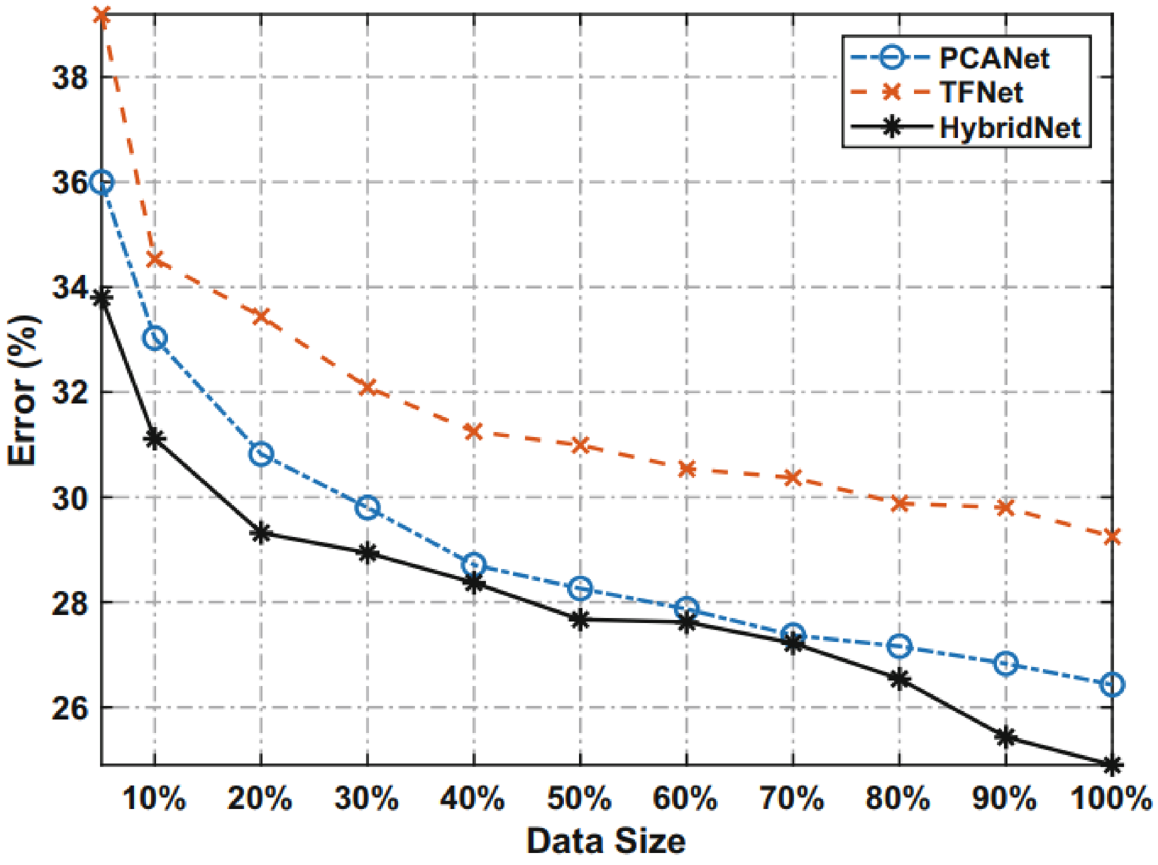
PCANet



CIFAR 10

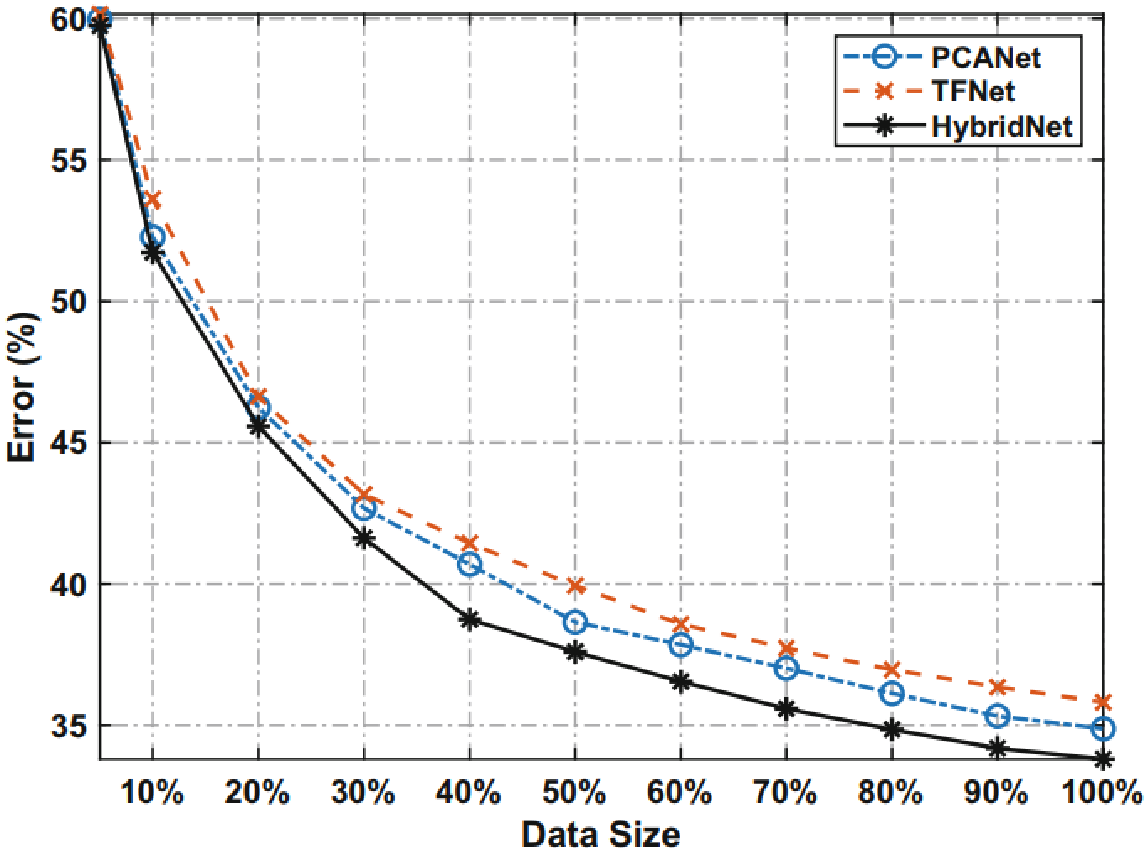
Results

PCANet



CIFAR 10

TensorNet



MNIST bg-img-rot

Contributions

- Contributions and findings
  - Both the amalgamated view and minutiae view of the data are individually insufficient.
  - To preserve the spatial information in the data TDNet is introduced.
  - HDNet is proposed to learn from both the views of the data.

Contributions, Conclusions

- Contributions and findings
  - Both the amalgamated view and minutiae view of the data are individually insufficient.
  - To preserve the spatial information in the data TDNet is introduced.
  - HDNet is proposed to learn from both the views of the data.

- Advantages and Limitations
  - Light Weight deep architecture which is unsupervised, fast, and less insensitive to noisy labels
  - Extremely challenging datasets might require more layers and non-linearities.

Contributions, Conclusions, and Future Works

- Contributions and findings
  - Both the amalgamated view and minutiae view of the data are individually insufficient.
  - To preserve the spatial information in the data TDNet is introduced.
  - HDNet is proposed to learn from both the views of the data.

- Advantages and Limitations
  - Light Weight deep architecture which is unsupervised, fast, and less insensitive to noisy labels
  - Extremely challenging datasets might require more layers and non-linearities.

- Future Works
  - Introduce fusion layer to combine the two networks.
  - Introduction mechanisms to handle rotations in the images.

THANK YOU

Sunny Verma