Team Ajinkya.Takawale

Ajinkya Takawale
Final year B.Tech, IIT Kharagpur
ajinkya.takawale97@gmail.com

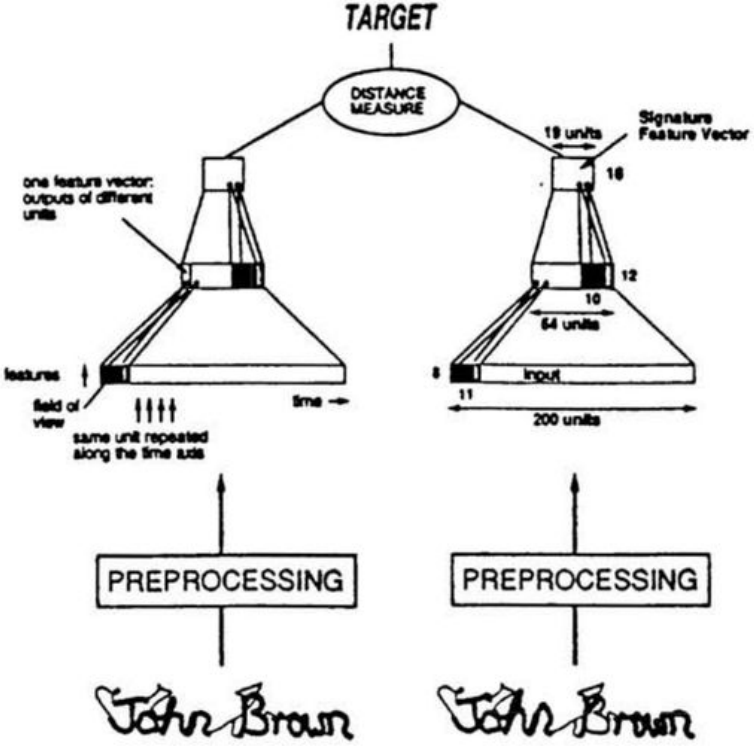
Linkedin **GitHub**

Sagar Panchal
Final year B.Tech, IIT Kharagpur
sgrpanchal31@gmail.com

Linkedin **GitHub**

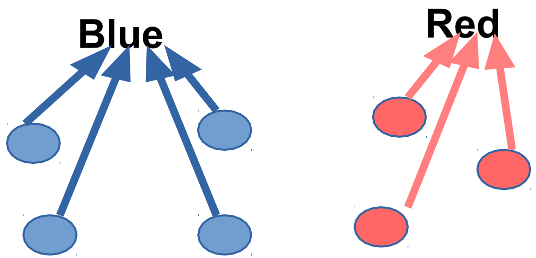
Siamese Networks

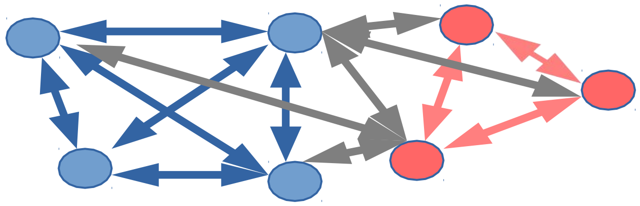
Signature Verification



Motivation

- Neural Networks can have unintentional behaviors.
  - Outliers
  - Model Bias
  - Unexplainable results
- Siamese Networks impose a structure that helps combat these problems.
- Siamese Networks allow us to use *more* data points than we would have in other common cases.
  - Allow us to use relationships between data points!

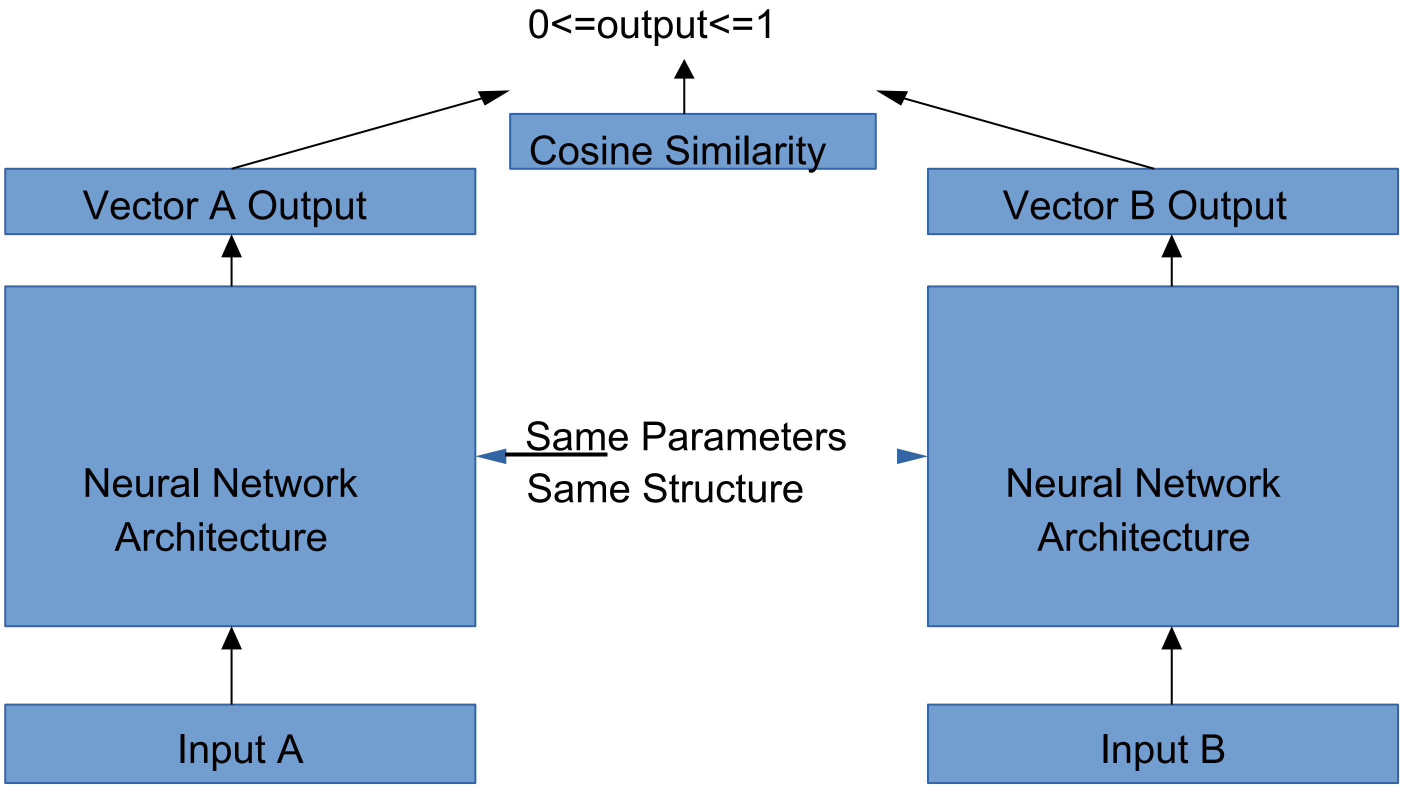




Structural Definition

- Siamese networks train a **similarity measure** between labeled points.
- Two input data points (textual embeddings, images, etc…) are run *simultaneously* through a neural network and are both mapped to a vector of shape Nx1.
- Then a standard numerical function can measure the distance between the vectors (e.g. the cosine distance).

Structural Definition



Training Dataset

- Siamese Networks must be trained on data that has two inputs and a target similarity.
  - ['input a1', 'input a2', 1]
  - ['input a2', 'input a3', 1]
  - ['input a2', 'input b1', 0]
  - …
- There must be similar inputs (0) and dissimilar inputs (1).

  Most studies have shown that the ratio of dissimilar to similar is optimal around:
  - Close to 1:1.
  - This depends on the problem and specificity of the model needed.

Training Dataset

- Since we have to generate similar and dissimilar pairs, the actual amount of training data is quite higher than normal.

- For example, in the initially given sample dataset there are around 300 total images and that gives us around 300C2 = 44850 unique pairs! So there should be no data shortage to train the network thus true to its name Siamese Networks are considered to be few-shot-learning models i.e being able to learn from less amount of data

- Also, I haven’t used any external dataset or additional dataset provided in later stages of competition

Dealing with New Data

- Another benefit is that siamese similarity networks generalize to inputs *and* outputs that have never been seen before.
  - This makes sense when comparing to how a person can make predictions on unseen instances and events.

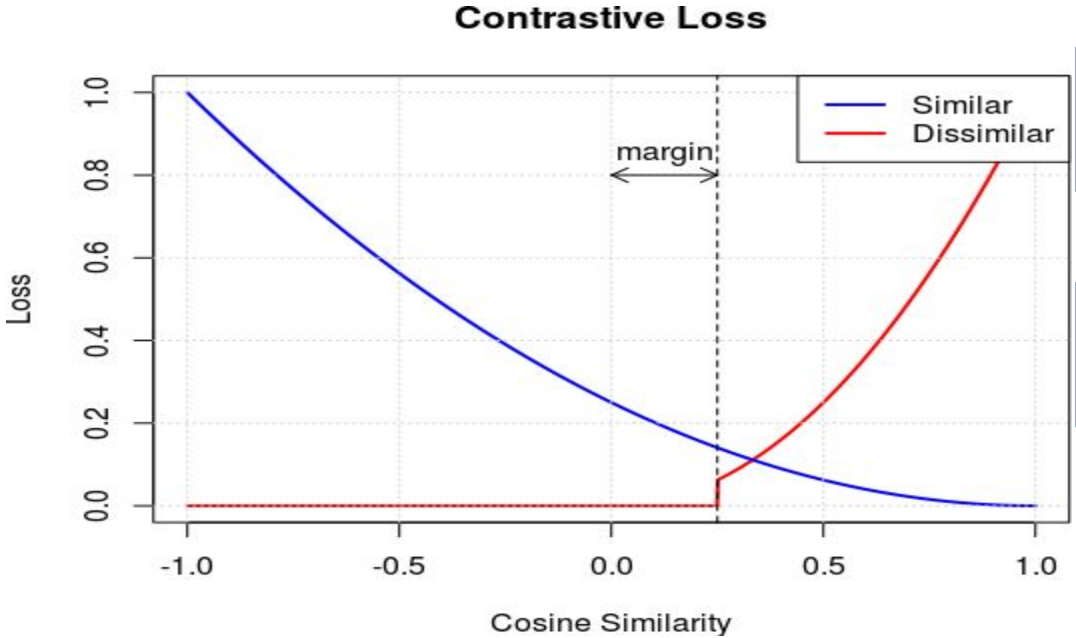
Help Explain Results

- With siamese networks, we can always list the nearest points in the output-vector space. Because of this, we can say that a data point has a specific label because it is nearest to a set of points.
- This type of explanation does *not* depend on how complicated the internal structure is.
- A possible application could be informing a user of potential forgery attempt by matching forged signature with closest matching genuine signature

Siamese Loss Function: Contrastive Loss

- The loss function is a combination of a similar-loss (L+) and dissimilar-loss (L-).

$$L = \sum (y_i \cdot L_+ + (1 - y_i) \cdot L_-)$$

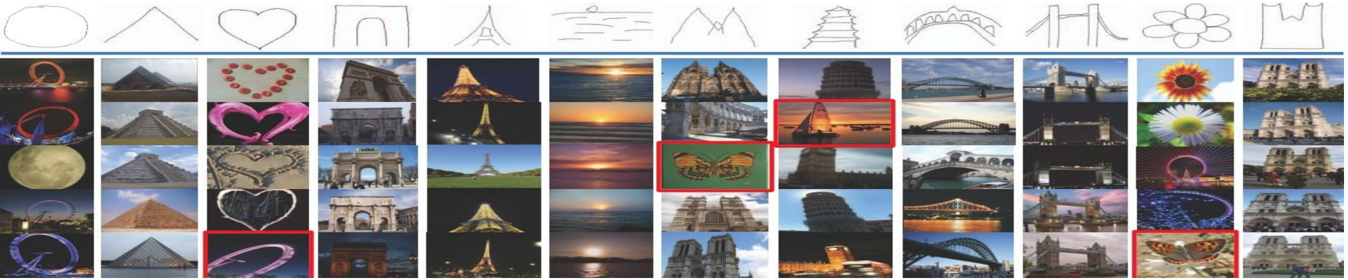


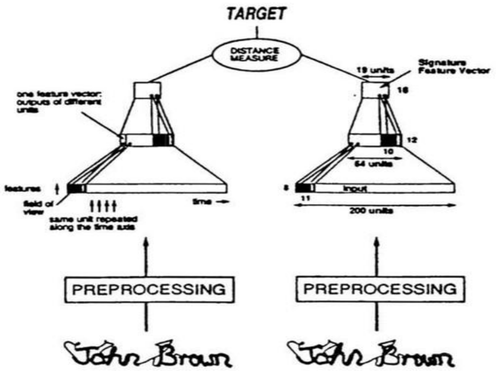
$$L_+ = \frac{1}{4} \cdot (1 - cosine(x_1, x_2))^2$$

$$L_- = \begin{cases} \text{cosine}(x_1, x_2)^2, & \text{if } x \geq \text{margin} \\ 0, & \text{otherwise} \end{cases}$$

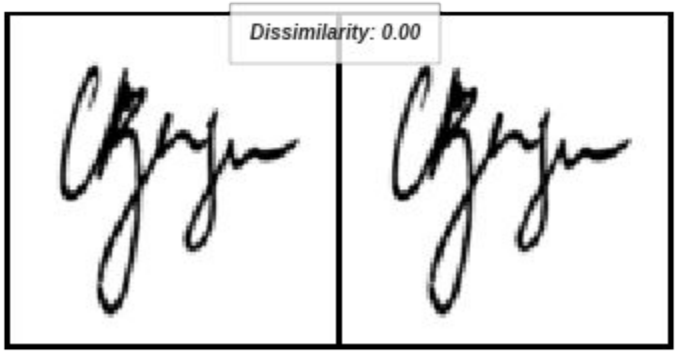
Potential Use Cases

- Natural Language Processing:
  - Ontology creation: How similar are words/phrases? Job Title
  - Matching: ‘VP of HR’ == ‘V.P. of People’
  - Topic matching: Which topic is this phrase referring to?
- Others:
  - Image recognition Image search
  - Signature/Speech recognition
  -

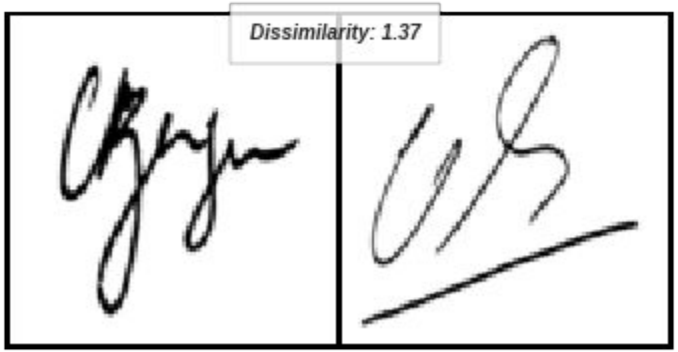




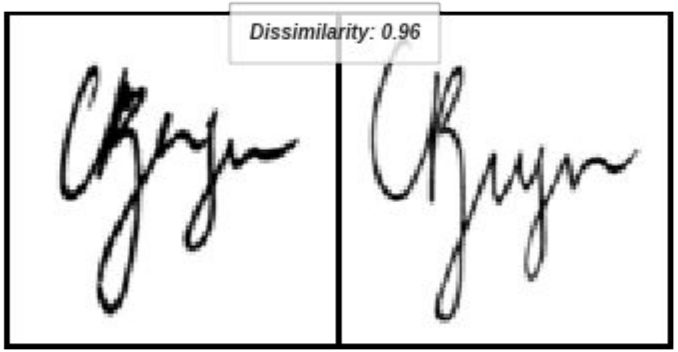
Results



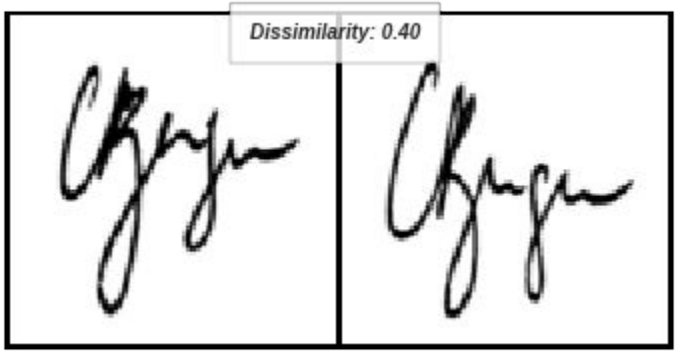
Left: Genuine Right: Genuine (same)



Left: Genuine Right: Different Person

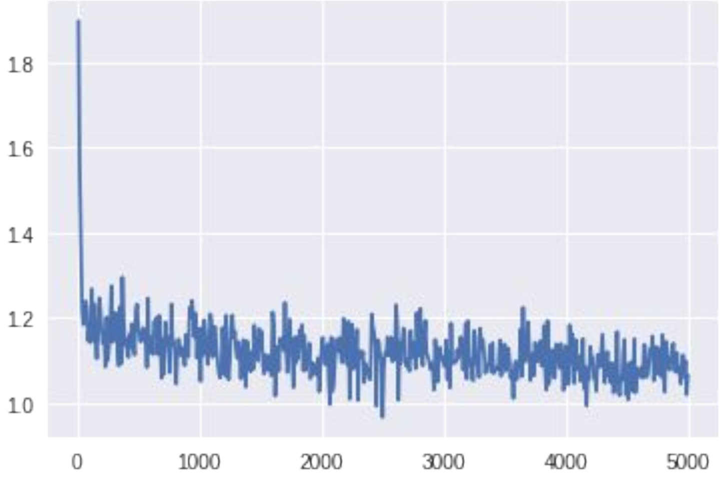


Left: Genuine Right: Forged

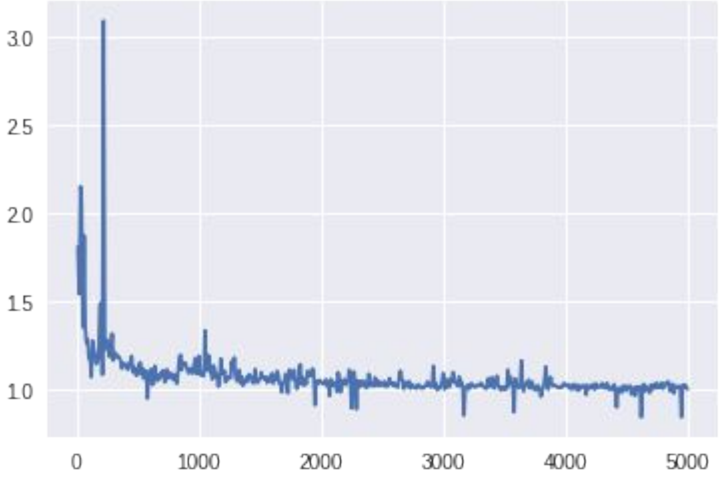


Left: Genuine Right: Genuine

Things I have learnt...



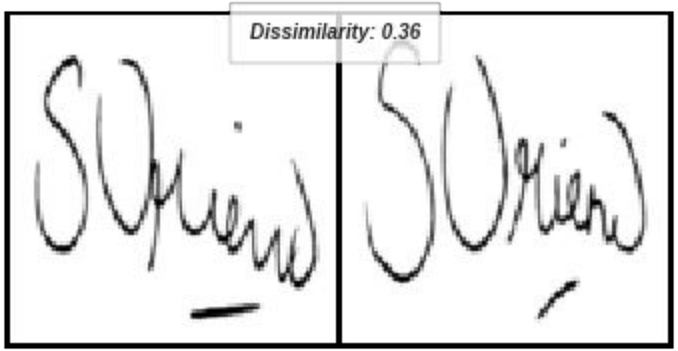
Case 1: Assuming two forged images signed by the same person as similar input to the network

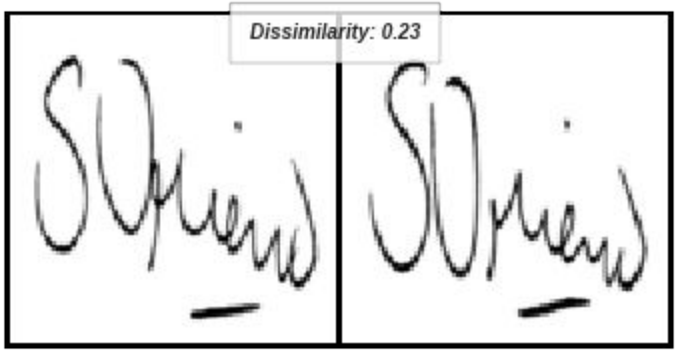


Case 2: Only considering a pair of genuine signatures as a similar input to the network

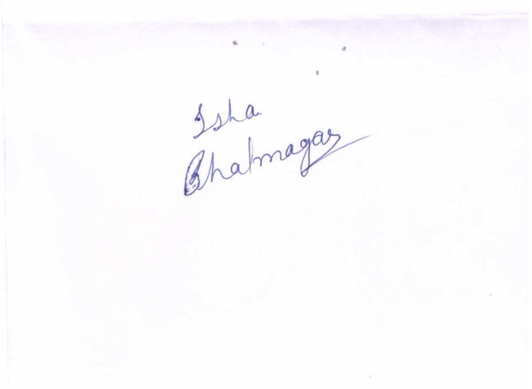
Conclusion: Loss is less fluctuating and has converged better in case 2. Also the hypothesis of considering two forged images signed by the same person as similar input to the network is false as it is really hard to replicate forged image in multiple attempts too!

Can you spot which one is forged and which one is genuine?





A few concerns regarding final dataset….



Conclusions and Summary

- Advantages:
  - Can predict out-of-training-set data.
  - Makes use of relationships, using more data.
  - Explainable results, regardless of network complexity.
- Disadvantages:
  - More computationally intensive (precomputation helps however).
  - More hyperparameters and fine-tuning necessary.
  - Generally, more training needed.
- When to use:
  - Want to exploit relationships between data points.
  - Can easily label ‘similar’ and ‘dissimilar’ points.

Further References

Signature Verification with fully connected siamese networks, 1995, Yann LeCun, et. al., Bell Labs,

http://papers.nips.cc/paper/769-signature-verification-using-a-siamese-time-

delay-neural-network.pdf

Attention based CNN for sentence similarity, 2015, https://arxiv.org/pdf/1512.05193v2.pdf

Learning text similarities with siamese RNNs, 2016, http://anthology.aclweb.org/W16-1617

Sketch-based Image Retrieval via Siamese CNNs, 2016, http://qugank.github.io/papers/ICIP16.pdf

Thank you for having fun!