Deep Learning approach to Handwriting Comparison

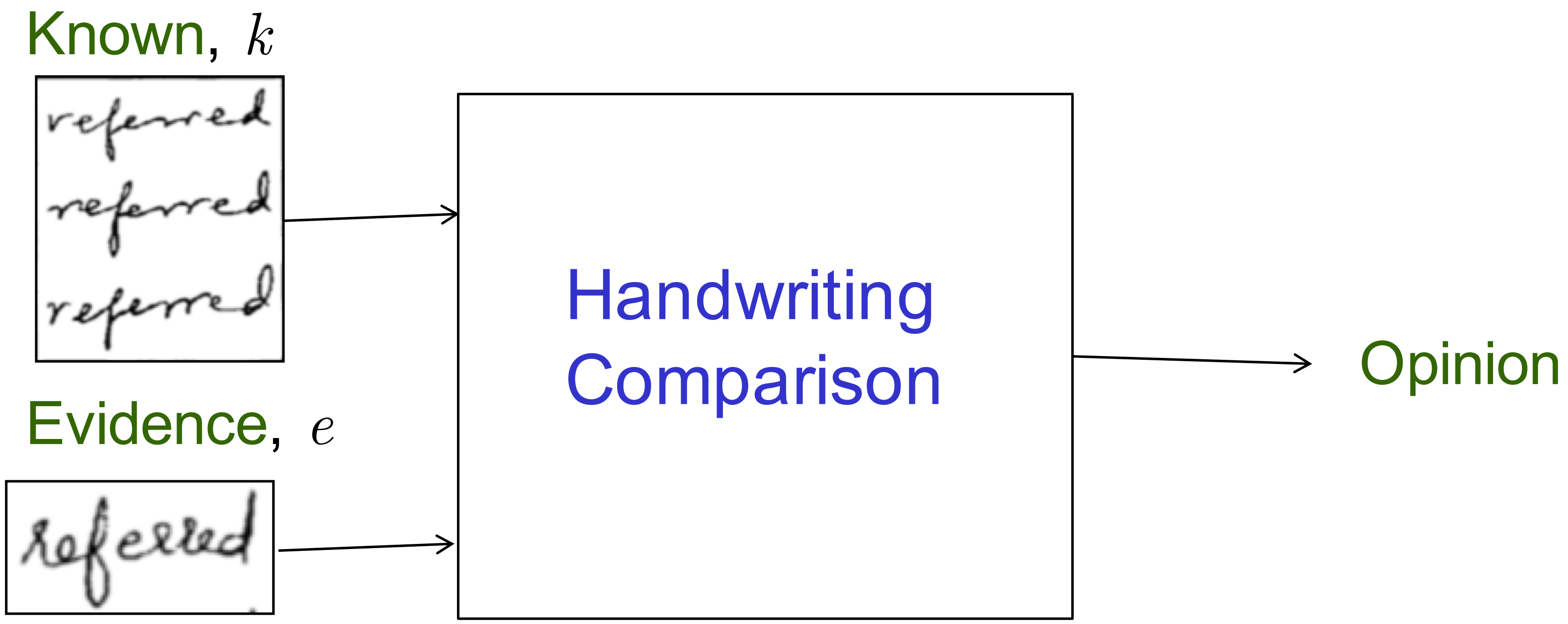
Sargur N. Srihari
University at Buffalo
The State University of New York
srihari@cedar.buffalo.edu

Plan of Presentation

1. Variability in Handwriting
2. The Handwriting Comparison Task
3. AI approaches to Handwriting Comparison:
   1. Knowledge-based Approach
   2. Classic Machine Learning Approach
   3. Deep Learning Approach
4. Four Deep Learning Architectures

Variability in handwriting

The Handwriting Comparison Task

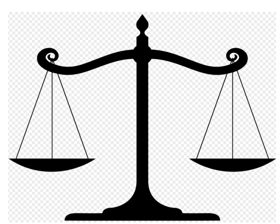


Opinion as a Likelihood Ratio

- h^0: Evidence (with characteristics) e was generated by *known* (individual with characteristics) k
- h^1: Evidence e was not generated by k

$$\text{Likelihood Ratio: } LR(k,e) = \frac{p(k,e \mid h^0)}{p(k,e \mid h^1)}$$

Probability of *evidence* and *known* (suspect) under Prosecution hypothesis



Probability of *evidence* and *known* (suspect) under Defense hypothesis

Opinion as a Probability

- Bayesian formulation
  - Define prior probabilities $p(h^0)$ and $p(h^1)$

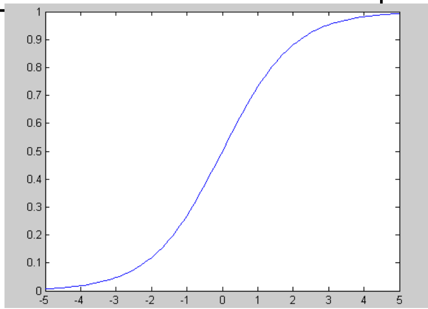
Define Prior Odds: $O_{prior} = \dfrac{p(h^0)}{p(h^1)}$ Then $p(h^0) = \dfrac{O_{prior}}{1+O_{prior}}$

Define $LR(k,e) = \dfrac{p(k,e \mid h^0)}{p(k,e \mid h^1)}$

Using Bayes rule, Posterior Odds: $O_{posterior} = O_{prior} \times LR(k,e)$ Then $p(h^0 \mid k,e) = \dfrac{O_{posterior}}{1+O_{posterior}}$

If $P(h^0) = P(h^1) = \dfrac{1}{2}$, $O_{prior} = 1$, $O_{posterior} = LR$

$$\boxed{p(h^0 \mid k,e) = \text{sigmoid}(LLR(k,e))}$$

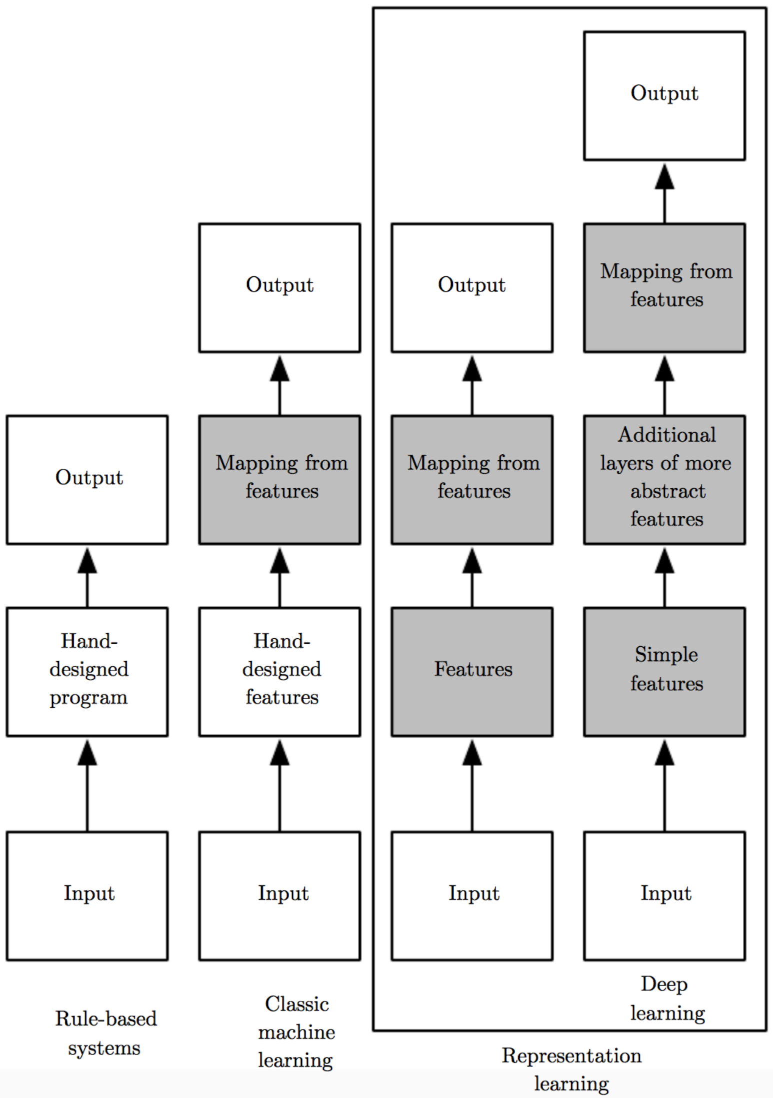


  - Can fuse results from different features
    - Additive LLRs under independence

Approaches to Arriving at Opinion

1. Knowledge (rule)-based approach
    1. Features/Rules are from FDE
2. Classic Machine Learning
    1. Features human engineered
    2. Decision is learned
3. Representation Learning
    1. Features/decision learned

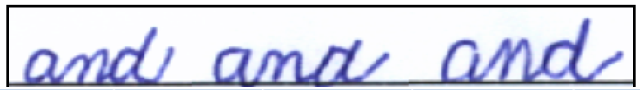
- Shaded boxes involve ML



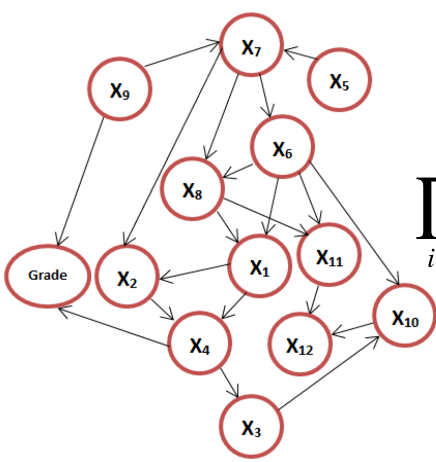
Approach 1: Knowledge-based

Features provided by FDEs

Cursive



Initial stroke of "a"	staff right	staff left	staff center	
Formation of "a" staff	tented	retraced	looped	no staff
Number of "n" arches	one	two		
Shape of "n" arches	pointed	rounded	retraced	combination
Location of "n" mid	above base	below base	at base	
Formation of "d" staff	tented	retraced	looped	
Formation of "d" initial	overhand	underhand	straight across	
Formation of "d" terminal	curved up	straight	curved down	no obvious end stroke
Symbol	unusual		symbol	
a-n relationship	a taller	a equal	a smaller	
a-d relationship	a taller	a equal	a smaller	
n-d relationship	n taller	n equal	n smaller	

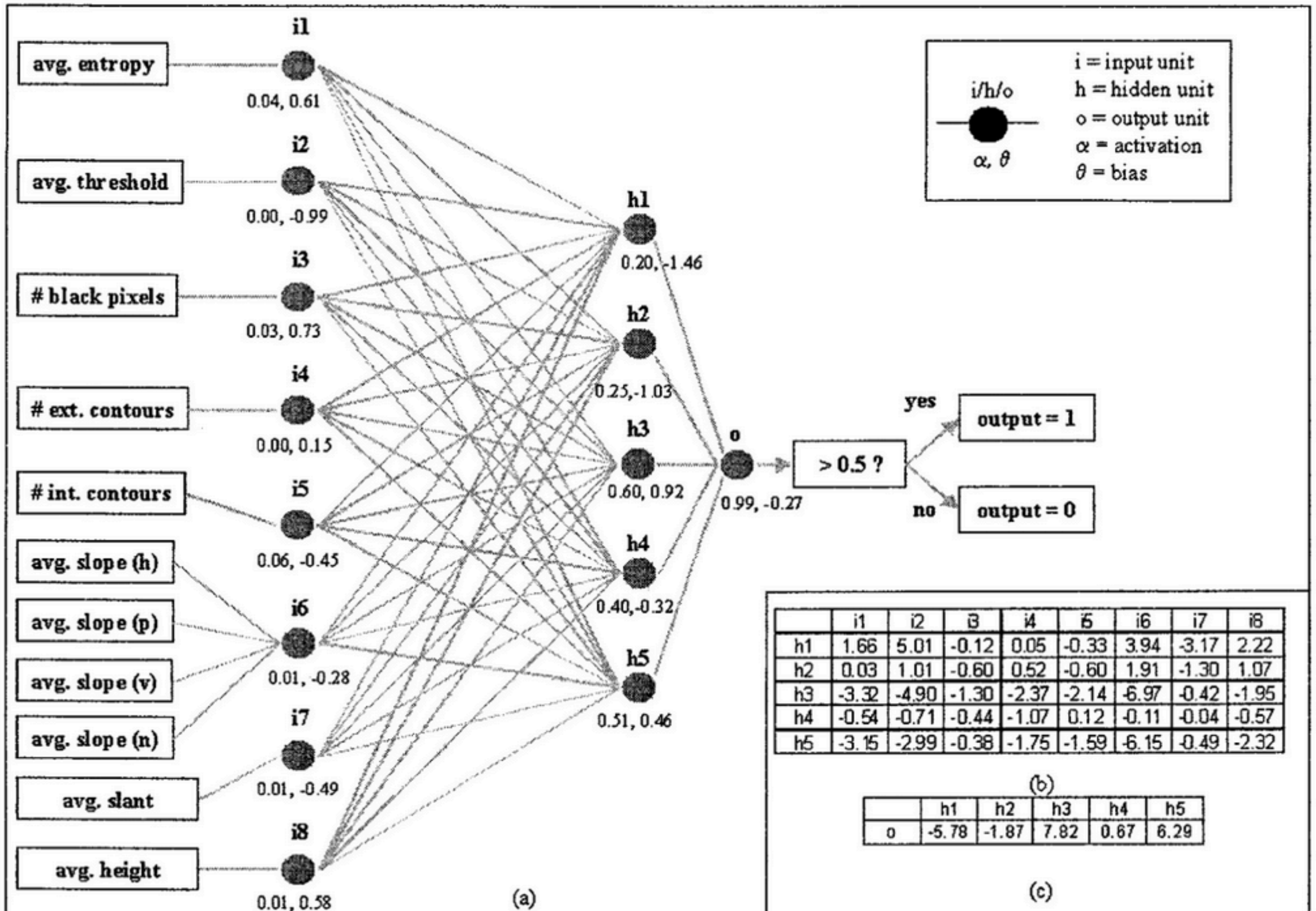


$$\prod_{i=1}^{n} p(x_i \mid pax_i)$$

Bayesian Network of feature differences

Results: Accuracy
Cursive : 75%

Approach 2: Simple machine learning



(a)

	i1	i2	i3	i4	i5	i6	i7	i8
h1	1.66	5.01	-0.12	0.05	-0.33	3.94	-3.17	2.22
h2	0.03	1.01	-0.60	0.52	-0.60	1.91	-1.30	1.07
h3	-3.32	-4.90	-1.30	-2.37	-2.14	-6.97	-0.42	-1.95
h4	-0.54	-0.71	-0.44	-1.07	0.12	-0.11	-0.04	-0.57
h5	-3.15	-2.99	-0.38	-1.75	-1.59	-6.15	-0.49	-2.32

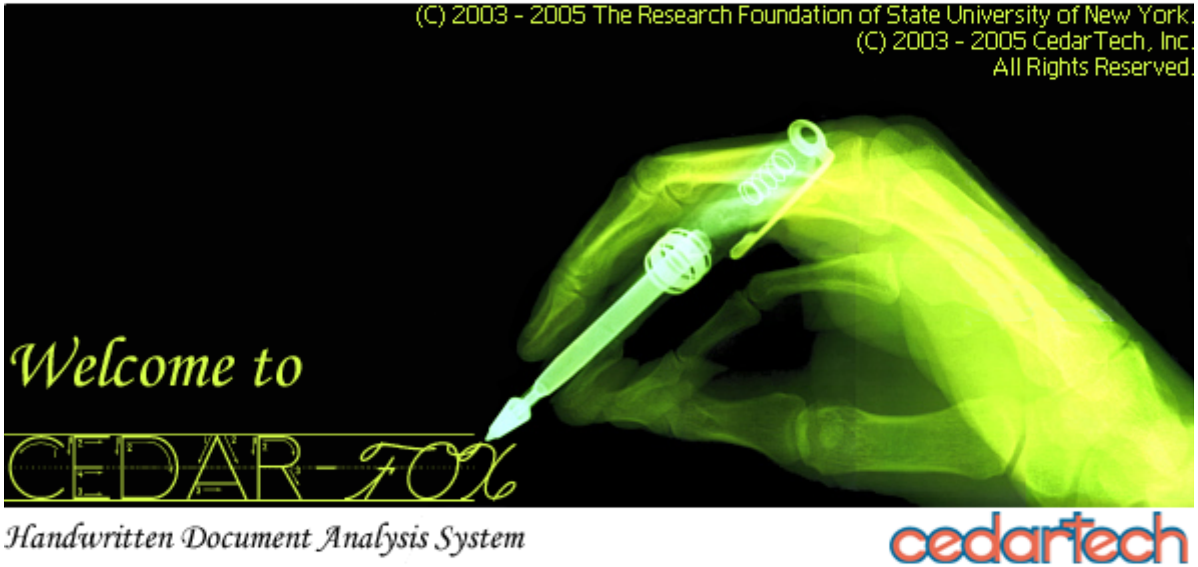
(b)

	h1	h2	h3	h4	h5
o	-5.78	-1.87	7.82	0.67	6.29

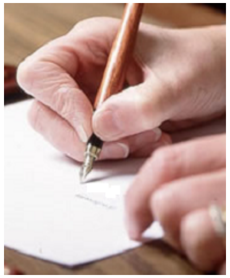
(c)

Approach 2: CEDARFOX software

- Writer Verification/Identification
  - Computes LLR based on human engineered features
  - Similarities based on Document Properties
    - Line Structure, Writer Characteristics



CEDAR-FOX: Side-by-side Comparison of Extended Writing



Known

Nov. 10, 1999

From
Jim Elder
829 Loop Street, Apt. 300
Allentown, New York 14707

To
Dr. Bob Grant
602 Queensberry Parkway
Omar, West Virginia 25638

We were referred to you by Xena Cohen at the University Medical Center. This is regarding my friend, Kate Zack.
It all started around six months ago while attending the "Rubeq" Jazz Concert. Organizing such an event is no picnic, and as President of the Alumni Association, a co-sponsor of the event, Kate was overworked. But she enjoyed her job, and did what was required of her with great zeal and enthusiasm.
However, the extra hours affected her health; halfway through the show she passed out. We rushed her to the hospital, and several questions, x-rays and blood tests later, were told it was just exhaustion.
Kate's been in very bad health since. Could you kindly take a look at the results and give us your opinion?

Thank You!

Jim

Questioned

November 10, 1999

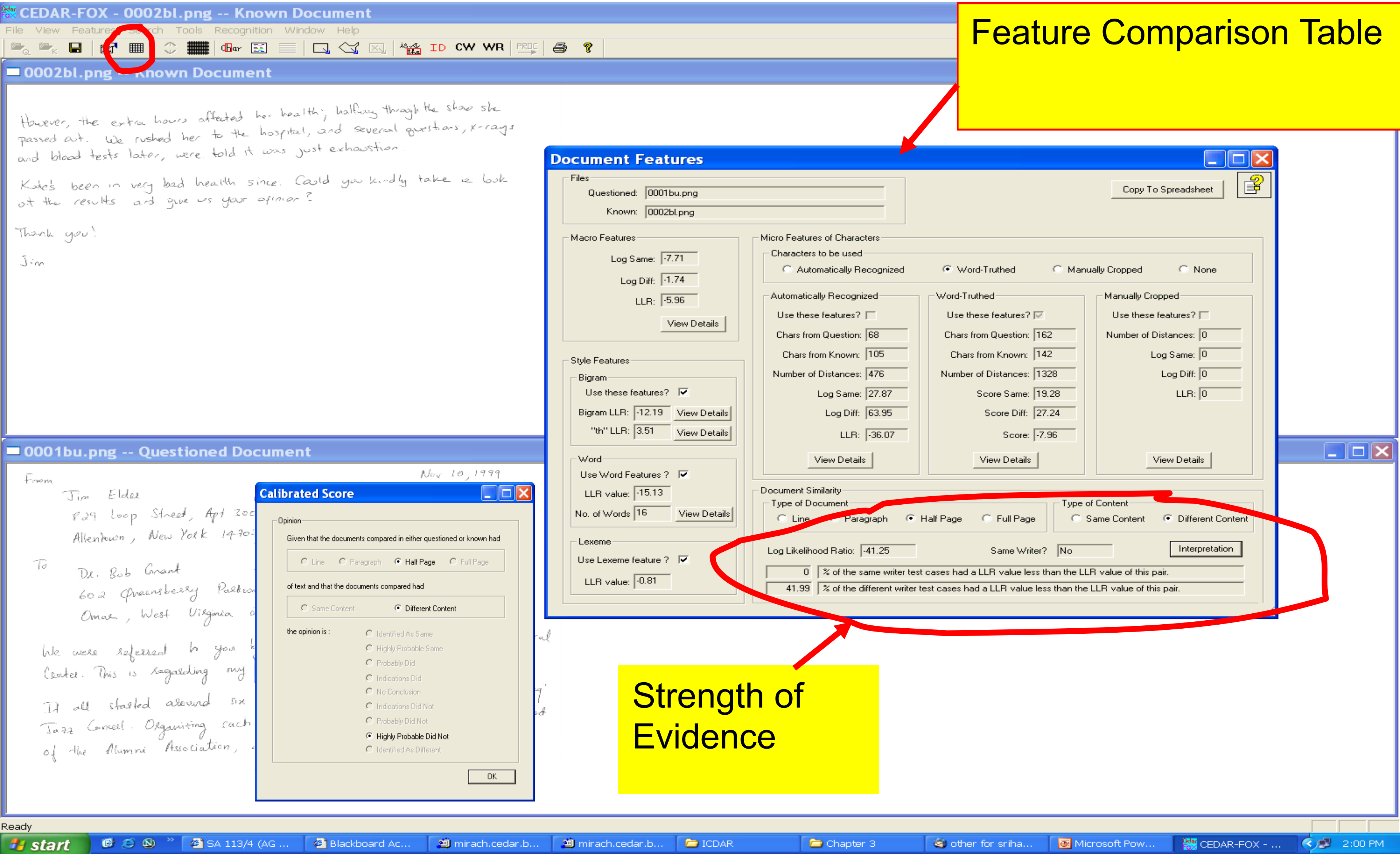
From
Jim Elder
829 Loop Street, Apt. 300
Allentown, New York 14707

To Dr. Bob Grant
602 Queensberry Parkway
Omar, West Virginia 25638

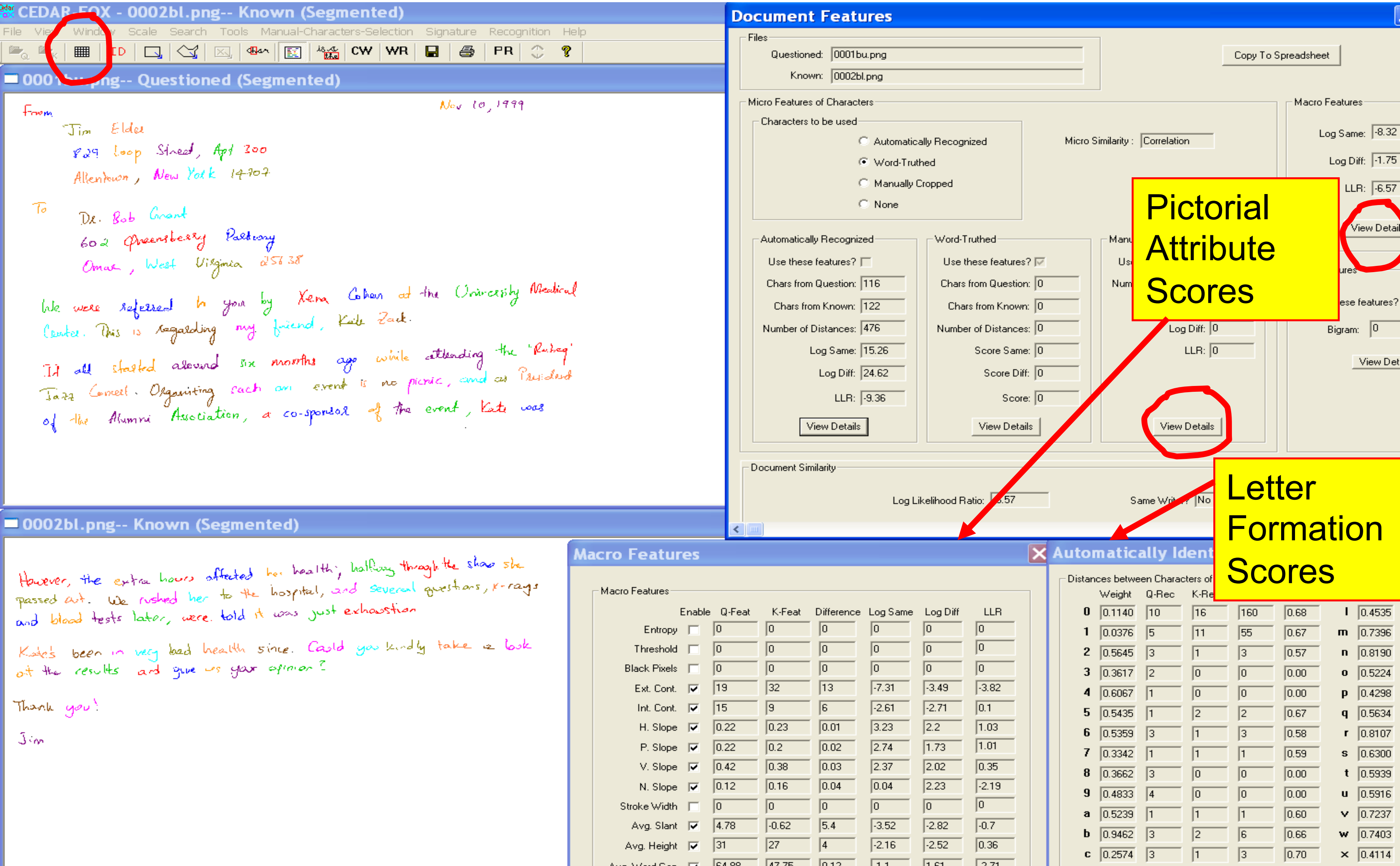
We were referred to you by Xena Cohen at the University Medical Center. This is regarding my friend, Kate Zack.

It all started around six months ago while attending the "Rubeq" Jazz Concert. Organizing such an event is no picnic, and as President of the Alumni Association, a cosponsor of the event, Kate was overworked. But she enjoyed her job, and did what was required of her with great zeal and enthusiasm.
However, the extra hours affected her health; halfway through the show she passed out. We rushed her to the hospital, and several questions, x-rays and blood tests later, were told it was just exhaustion.

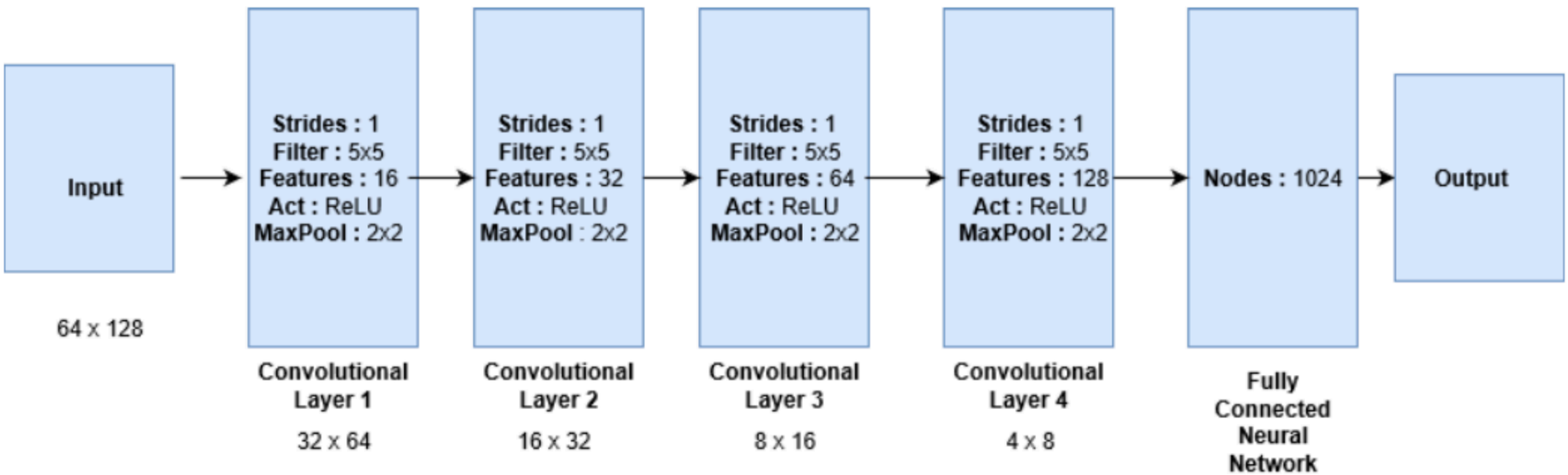
Result of Comparsion



Human Engineered Characteristics



Approach 3: Deep Learning Architecture 1



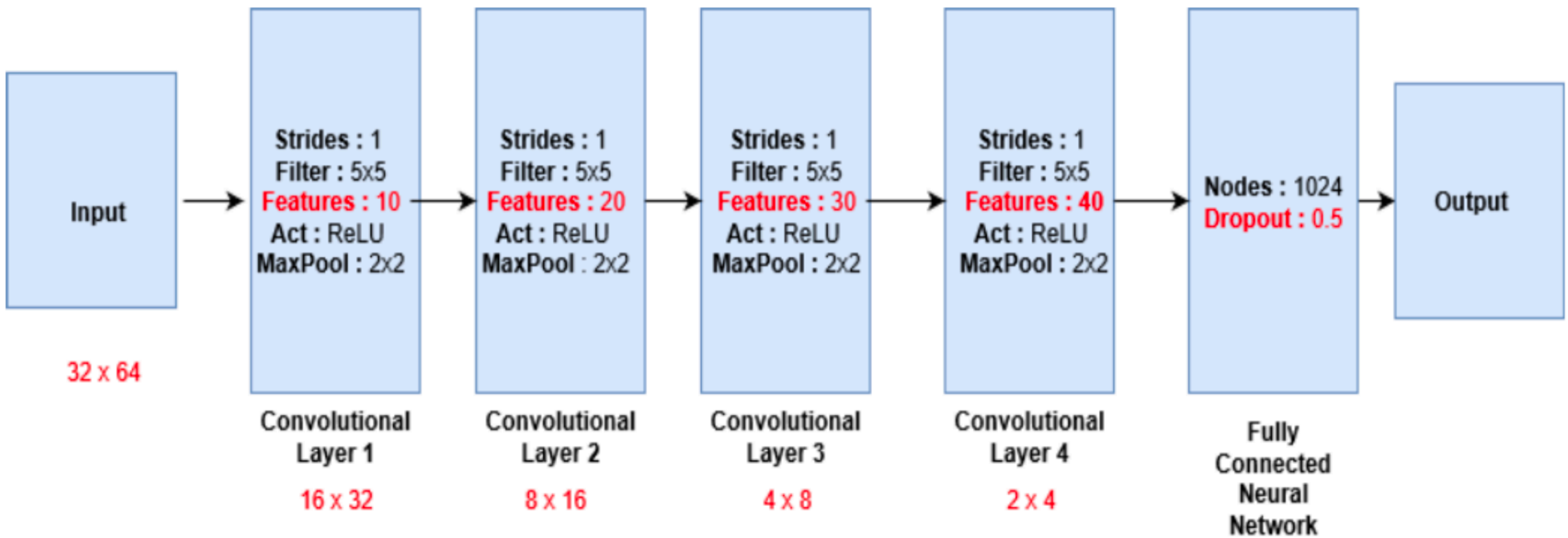
Stack images and feed to a 4 layer CNN (feature extractor) and a FCN for classification

Results

Type -1 : 59%

Type 2 : 65%

Deep Learning Architecture 2



First resize images and standardize on image padding and strides

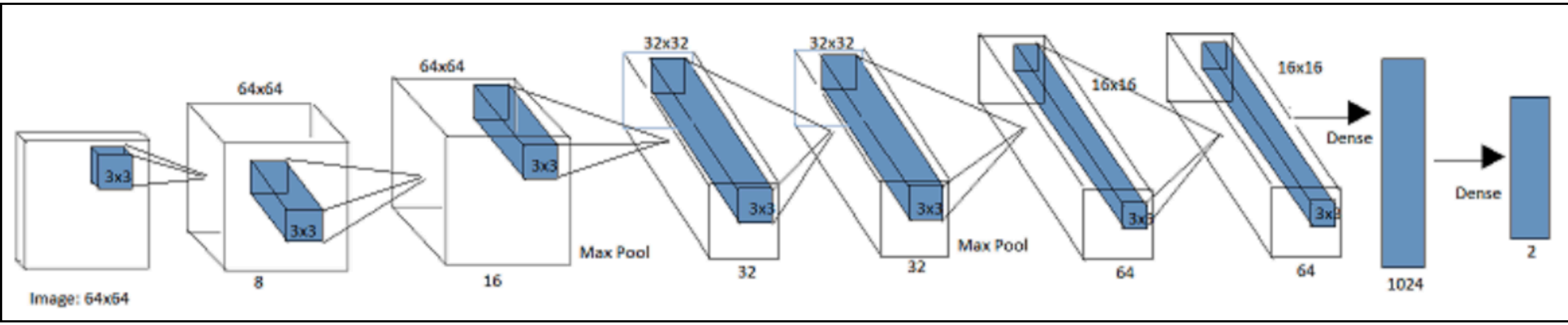
Include regularization: dropout in the fully connected layer

Results

Type -1 : 85%

Type -2 : 78%

Deep Learning Architecture 3



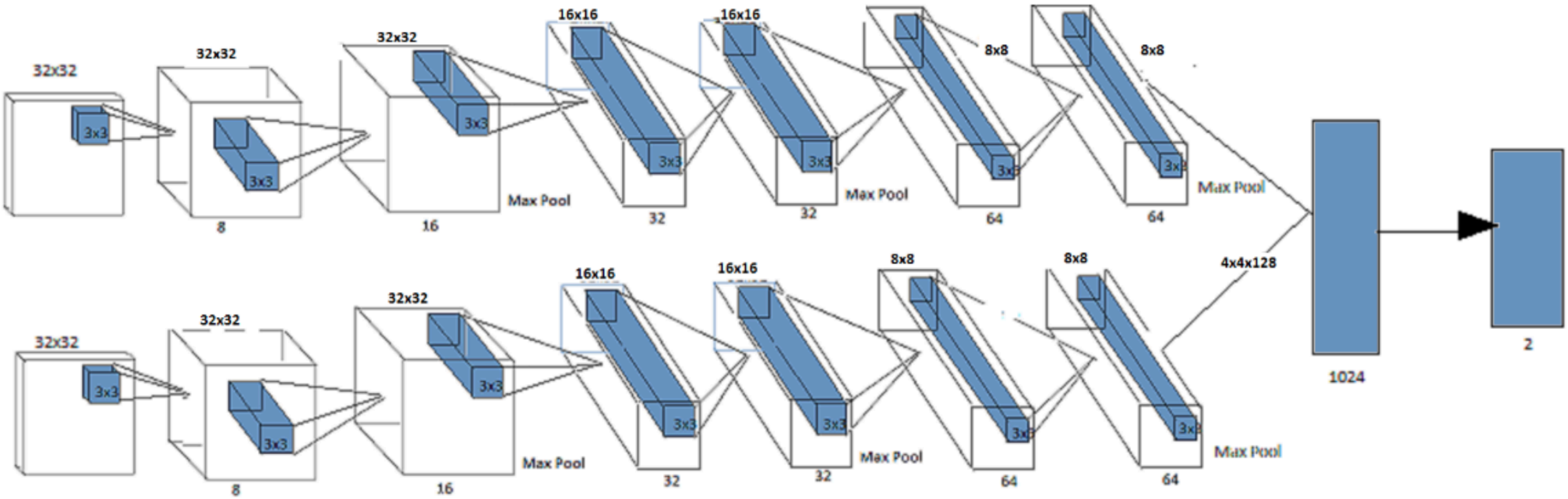
Introduce maxpooling and batch normalization

Results:

Type -1 : 95%

Type -2 : 84%

Deep Learning Architecture 4



Using same architecture as above but split into two channels

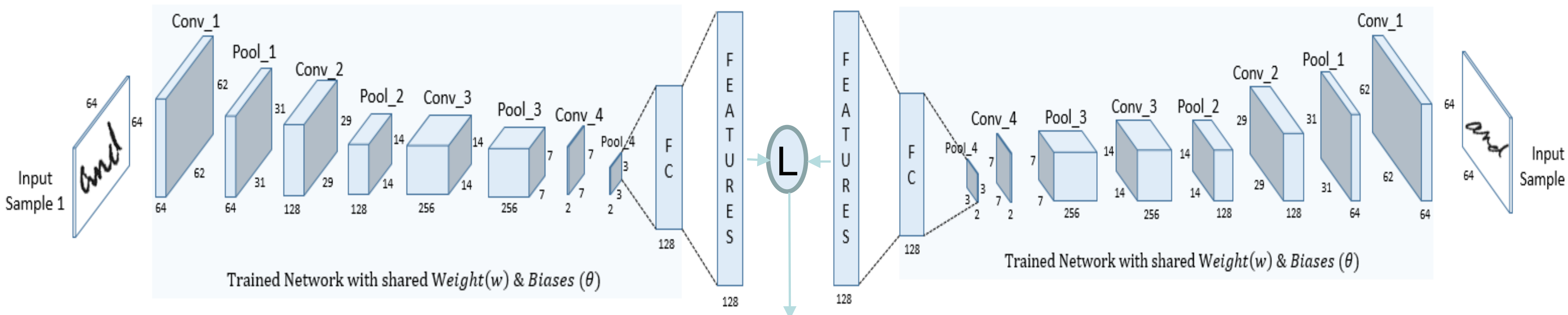
Results:

Type -1 : 96%

Type -2 : 89%

Architecture 5: One Shot Twin Network

https://github.com/mshaikh2/DeepLearningHandwritingForensics



Log likelihood Contrastive Similarity

$ln[C_{Loss}(M_{w-sim})/C_{Loss}(M_{w-dis})]$

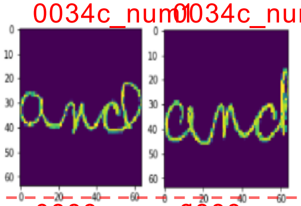
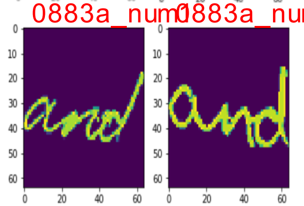
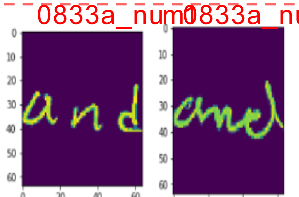
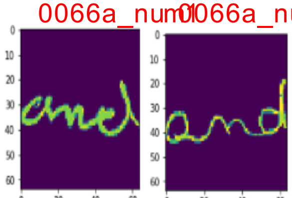
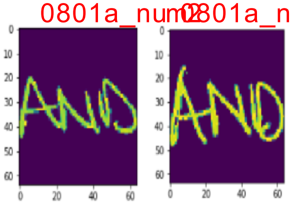
$C_{Loss}(M_{w-dis}) = 0.5 * \{max(0, m - D_w)\}^2$

$C_{Loss}(M_{w-sim}) = 0.5 * \{D_w\}^2$

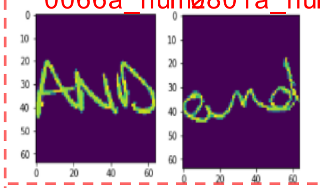
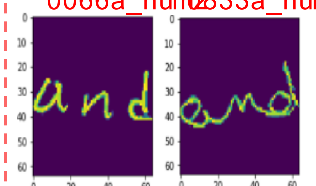
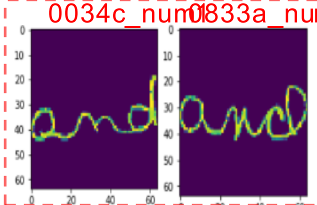
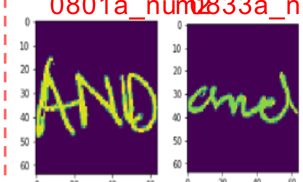
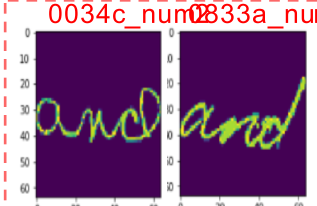
Algorithm 1: One Shot Twin Network
Data: Known & Query Image
Result: Writer Verification
Known Image: X_i & Query Image: X_j ;
Dissimilarity Model: M_{w_diss}; Weights of M_{w_diss}: W_dis
Similarity Model: M_{w_sim}; Weights of M_{w_sim}: W_sim_{ij} ;
Contrastive Loss: C_{Loss};
for X_i .. X_j **do**
  Update W_diss_{ij} for M_{w_diss} assuming samples classifying in different classes [0,1] ;
  Update W_sim_{ij} for M_{w_sim} assuming samples classifying in the same classes [1] ;
end
for *Model in* $[M_{w\_diss}, M_{w\_sim}]$ **do**
  for *Datapoint in* $[X_i .. X_j]$ **do**
    Extract features $F_{Datapoint}[0..127]$ for Datapoint from Model;
  end
  Calculate: C_{Loss} $(F_i[0..127], F_j[0..127])$
end

Comparison of Two Approaches

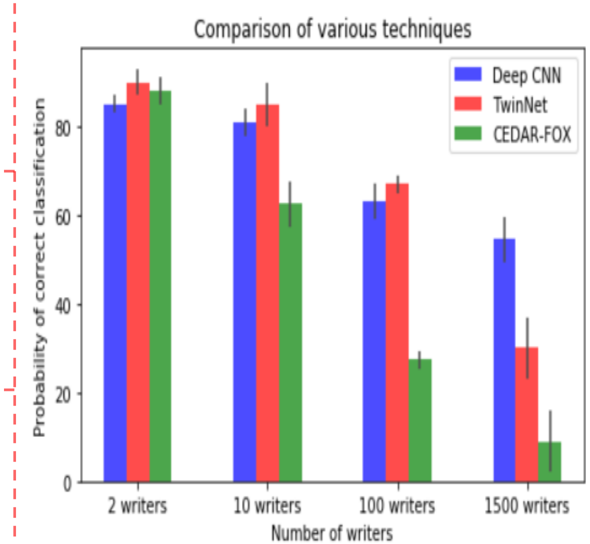
TYPE I: Intra Class Accuracy (LLR)

Known Quest	CEDAR-FOX	OSTN
0034c_num0034c_num2 	-19. ✗	0.308 ✓
0883a_num0883a_num3 	-13.24 ✗	-0.59 ✗
0833a_num0833a_num2 	-22.45 ✗	0.054 ✓
0066a_num0066a_num2 	-5.84 ✗	0.76 ✓
0801a_num0801a_num2 	8.92 ✓	0.19 ✓

TYPE II: Inter Class Accuracy (LLR)

Known Quest	CEDAR-FOX	OSTN
0066a_num2801a_num1 	-33. ✓	-0.62 ✓
0066a_num0833a_num2 	-16. ✓	-0.707 ✓
0034c_num0833a_num1 	-27. ✓	-1.134 ✓
0801a_num0833a_num1 	-25.9 ✓	-1.08 ✓
0034c_num0833a_num1 	-30.65 ✓	-1.06 ✓

Performance with no of writers



Comparison	Type I	Type 2
CEDAR-FOX	77%	88%
OSTN	90%	96%

Negative values measure dissimilarity, Positive

Summary

- Compared performance of three AI approaches for the task of handwriting comparison
  1. Knowledge-based: Features specified by FDEs
  2. Simple machine learning: Features engineered
  3. Deep learning: Features learnt automatically
- Observed progressive performance improvement over the three approaches