CAPTCHAs

• Computer generated test that distinguishes humans from computers
  – Relies on gap between human and computer abilities

• Usually a low-level sensory task
  – “Identify these [really distorted] digits”

• Used to prevent bot-based abuse
Create your Hotmail account

This is your Windows Live ID—it gets you into other services like Messenger and SkyDrive. All information is required.

If you use Hotmail, Messenger, or Xbox LIVE, you already have a Windows Live ID. Sign in

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hotmail address</td>
<td>@ hotmail.com</td>
</tr>
<tr>
<td>Create a password</td>
<td>5-character minimum; case sensitive</td>
</tr>
<tr>
<td>Retype password</td>
<td></td>
</tr>
<tr>
<td>Mobile phone</td>
<td>United States (+1)</td>
</tr>
<tr>
<td>Alternate email address</td>
<td>Or choose a security question for password reset</td>
</tr>
<tr>
<td>First name</td>
<td></td>
</tr>
<tr>
<td>Last name</td>
<td></td>
</tr>
<tr>
<td>Country/region</td>
<td>United States</td>
</tr>
<tr>
<td>ZIP code</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Male or Female</td>
</tr>
<tr>
<td>Birth date</td>
<td>Month, Day, Year</td>
</tr>
</tbody>
</table>

Enter the characters you see

New | Audio | Help

Send me email with promotional offers and survey invitations from Windows Live, Bing, and MSN. (You can unsubscribe at any time.)
Gmail
A Google approach to email.

Gmail is built on the idea that email can be more intuitive, efficient, and useful. And maybe even fun. After all, Gmail has:

- **Lots of space**
  Over 7695.415227 megabytes (and counting) of free storage.

- **Less spam**
  Keep unwanted messages out of your inbox.

- **Mobile access**
  Get Gmail on your mobile phone. [Learn more](#)

[Sign in to Gmail](#)

**Enter the letters above**

Letters are not case-sensitive

[Sign in] [Stay signed in]

[Can't access your account?](#)

[Take Gmail to work with Google Apps for Business](#)

Love Gmail, but looking for a custom email address for your company? Get business email, calendar, and online docs @your_company.com. [Learn more](#)
Machine Learning Attacks

Input
Machine Learning Attacks

Input

Preprocessing

p v a c t
Machine Learning Attacks

Input, Preprocessing

Segmentation

PRACT
Machine Learning Attacks

Input, Preprocessing, Segmentation

Postprocessing
Machine Learning Attacks

Input, Preprocessing, Segmentation, Postprocessing

Classification

pvack
Training

- Preprocessing
- Segmentation
- Postprocessing

Unsupervised, Specific to schema

Classification

1. Solve CAPTCHAs (Amazon Turk, grad students)
2. Run CAPTCHAs through unsupervised stage
3. Train
Audio CAPTCHAs

- Input

Audio CAPTCHAs

• Segmentation
Audio CAPTCHAs

- Postprocessing: Cepstrum representation
- Classification: Regularized Least Squares

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authorize</td>
<td>87.25</td>
</tr>
<tr>
<td>Digg</td>
<td>40.84</td>
</tr>
<tr>
<td>eBay</td>
<td>82.88</td>
</tr>
<tr>
<td>Microsoft</td>
<td>48.95</td>
</tr>
<tr>
<td>Recaptha</td>
<td>1.52</td>
</tr>
<tr>
<td>Yahoo</td>
<td>45.45</td>
</tr>
</tbody>
</table>
Visual CAPTCHAs

• Preprocessing & Segmentation

Visual CAPTCHAs

- Preprocessing & Segmentation

k b j r z 9
Visual CAPTCHAs

• Preprocessing & Segmentation

k b j r z g  75%
Visual CAPTCHAs

- Preprocessing & Segmentation

75%

nuexdp
Visual CAPTCHA

- Preprocessing & Segmentation

75%
Visual CAPTCHAs

- Preprocessing & Segmentation

75% 30%

nuce5p
Visual CAPTCHAs

• Preprocessing & Segmentation

[k b i r z g] 75%

[nucx5p] 30%

[imcvbp]
Visual CAPTCHAs

- Preprocessing & Segmentation

75% 30%
Visual CAPTCHAs

- Preprocessing & Segmentation

- 75%
- 30%
Visual CAPTCHAs

- Preprocessing & Segmentation

  - Image 1: 75%
  - Image 2: 30%
Visual CAPTCHAs

• Preprocessing & Segmentation

- 75%
- 30%
- 96%
Visual CAPTCHAs

- Preprocessing & Segmentation

- $75\%$
- $96\%$
- $30\%$
Visual CAPTCHA As

• Preprocessing & Segmentation

75%  30%  96%
Visual CAPTCHAs

- Preprocessing & Segmentation

[Images with text]
- Image 1: kbjrzg with 75%
- Image 2: nucx5p with 30%
- Image 3: lmcvbp with 96%
Visual CAPTCHAs

- Preprocessing & Segmentation

- Preprocessing: 75%
- Segmentation: 86%
- Recognition: 30%
- Noise: 96%
Visual CAPTCHAs

- Preprocessing & Segmentation

- k bj rz g: 75%
- p vac f: 86%
- nucx5p: 30%
- dissents: 96%
- Jmcvbp: 96%
Visual CAPTCHAs

• Preprocessing & Segmentation

75% 30% 96%
Visual CAPTCHAs

• Preprocessing & Segmentation

- 75% for 'kbjrzg'
- 30% for 'nucox5p'
- 96% for 'jmcvdp'
- 86% for 'pvacl'
- 52% for 'dissents'
Visual CAPTCHAs

- Preprocessing & Segmentation

- 75%
- 86%
- 30%
- 52%
- 96%
- 0%
Common Weaknesses

• Segmentation
  – Audio: background noise sounds like digits
  – Visual: collapsing prevents extraction of digits

• Segmentation is separate from Classification
  – Easier to train
  – Segmentation is weak classification

• Must create new breaker for each scheme
  – Often requires hand tuning!
A Different Paradigm

• How do people do it?
• Training
  – Entire lifetime of seeing/hearing warped letters
  – Single breaker
• Solving
  – Naturally invariant to transformations
  – Neocortex is a “deep” architecture
  – Segmentation and Classification are the same
Deep Learning

• Give a very brief overview of deep learning in that it processes input in alternating layers (template matching & max pooling)

• Discuss how training these models leads them to learn how to represent data so that they can recognize digits. For instance, I would like the computer to infer that it should perform a cepstrum transform (and in fact there is a paper that shows exactly this using deep learning). What’s also cool is that if you run this on visual input, you get templates that match the processing that goes on in the cortex – maybe we can actually get human-like performance with this.

• In the way of invariance, explain how abstraction increases as we go up the hierarchy of the deep learning and that each layer is invariant in its own way and this leads to a very complicated invariance that may help explain why humans are so invariant to the distortions.
Deep Learning

Abstraction, Stability

Complex

Simple

Complex

Simple
Invariance Manifolds

• The particulars of what goes on in each template matching/ max pooling pair of layers is what leads to different deep learners (name a few).

• We’re considering ones that focus on invariance. Inside each pair of layers, we will be using algorithms from a paper that I’m currently submitting that uses the idea of invariance manifolds.

• Define what invariance manifolds are.
Invariance Manifolds

- Objects can’t be represented as single points in pixel space
- Represent as manifold $f: \mathbb{R}^p \rightarrow \mathcal{P}$ parameterized by possible configurations
Looking Ahead

• How will deep learning compare to hand-tuned crackers?

• How much training will it take to handle multiple schemes?
  – Measure of CAPTCHA difficulty?

• What are the invariance properties of these networks?
  – Beneficial to AI and security
Looking Ahead

• Performance advantages over hand-tuned methods
• Single algorithm to benchmark CAPTCHAs
  – Several measures of scheme’s difficulty
• Characterization of invariance properties and weaknesses
• Better understanding of CAPTCHA design