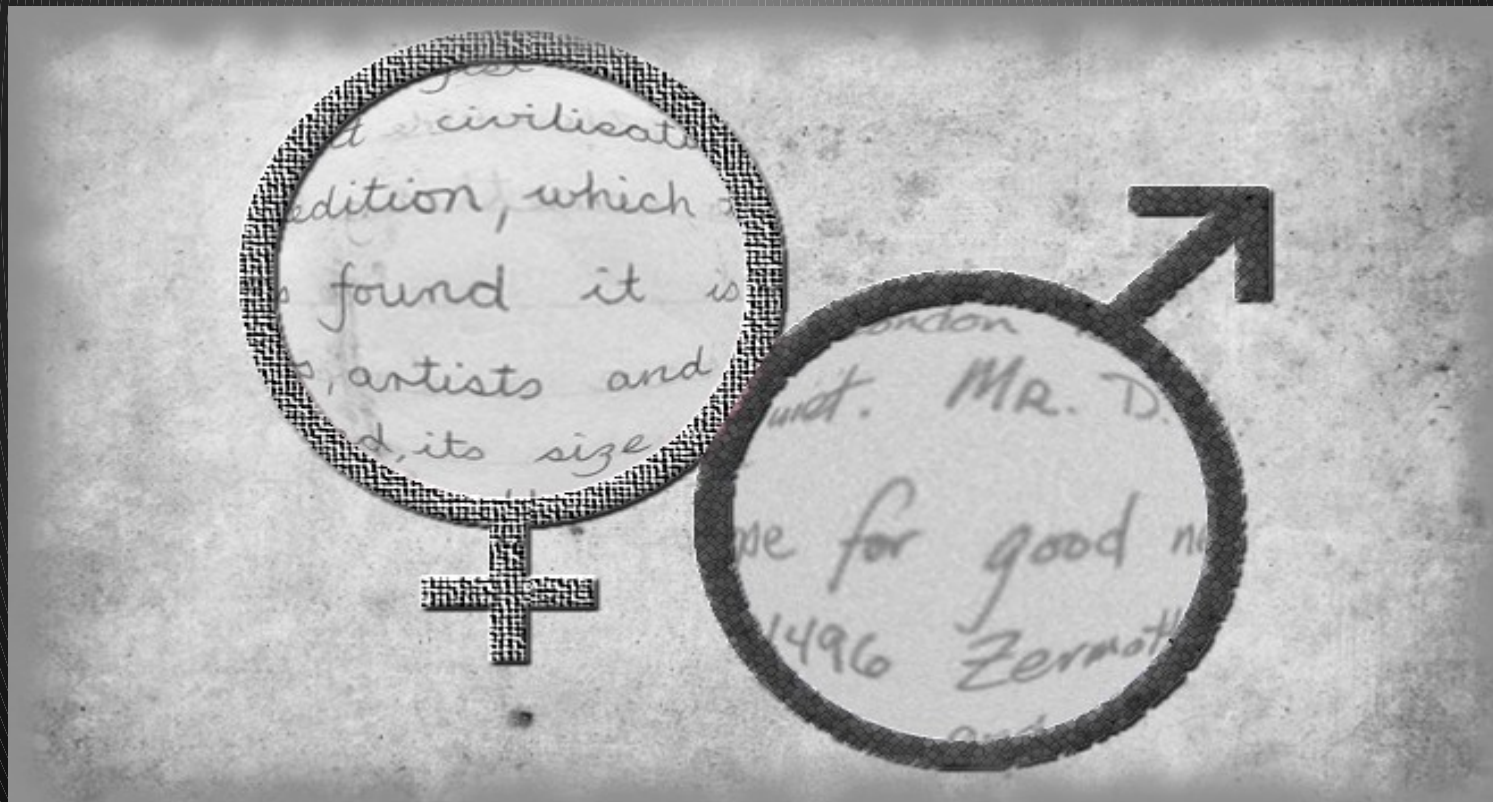


# Gender Prediction from Handwriting (ICDAR 2013)

Aakash Gupta, Utsav Bansal



# Dataset Images-282 writers

\* فعل الأمر فعل يطلب به إلى المخاطب القيام بعمل معين في المستقبل .  
\* يبني فعل الأمر في الأهل على السكون الظاهر على آخره .  
\* إذا كان فعل الأمر معتلاً الآخر يبني على حذف حرف العلة من آخره .  
\* إذا صيغ فعل من الأفعال الخمسة يبني على حذف التون من آخره .  
\* إذا كان فعل الأمر مهناعف الآخر يبني على السكون و يترك بالفتح منعاً من التقاء الساكنين . بدأ القفل رزق الله يخبو منذ الشهر الثامن .

الزلازل هو ظاهرة طبيعية عبارة عن اهتزاز أرضي سريع يعود إلى تكسر الصخور وإزاحتها بسبب تراكم إجهادات داخلية نتيجة لهزات هبلوجية ينجم عنها تحرك الصنائح الأرضية. وتكون أحياناً آثاره كارثية على أي منطقة تتعرض له.  
قد ينشأ الزلازل كنتيجة لأنشطة البراكين أو نتيجة لوجود انزلاقات في طبقات الأرض. وتؤدي الزلازل إلى تشقق الأرض ونهب الينابيع أو ظهور الينابيع الجديدة أو حدوث أعواج عالية لها آثارها التخريبية للمباني والمواصلات والمنشآت.  
وعالماً ينتج عن حركات الحمل الحراري في الإستينوسفير والتي تحرك الصنائح القارية متسببة في حدوث هزات هي الزلازل. إذ أن مختلف المناطق المرجوة على خط الزلازل تأخذ دوراً التدابير الاحترازية.

Arabic Page with different text

Arabic Page with same text

Claire holds up her hands to show her nails. "Don't they look awful!" she cries. "I so miss New York manicures." That's not all Claire misses about her other life in the USA. For two years she has divided her life between her home town, Manchester, in the north of England, and Brooklyn, New York. She has an art gallery and a home in both places. She flies once a month, and spends about a month at a time in each. "I dress differently in New York

The International Organization for Migration (IOM) said there are more than 200 million migrants around the world today. Europe hosted the largest number of immigrants, with 70.6 million people in 2005, the latest year for which figures are available.

North America, with over 45.1 million immigrants, is second, followed by Asia which hosts nearly 25.3 million. Most of today's migrant workers come from Asia. The United Nations estimates that there are 814 million migrants across the globe, an increase of about 37% in two decades.

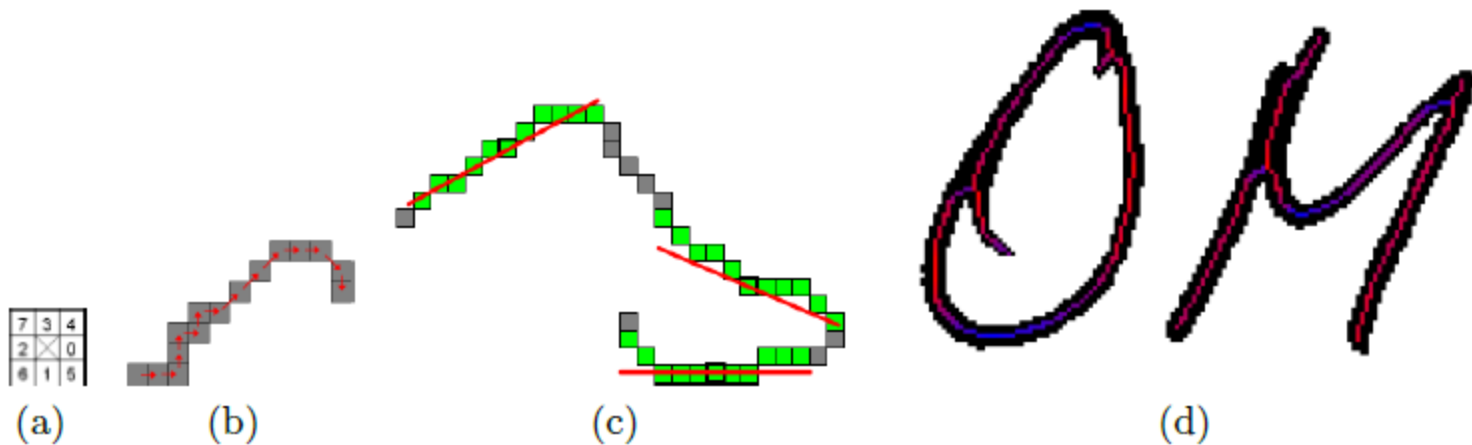
Also the immigration is not only destined to America and Europe, but we can see a large amount of immigration from Asia countries to the Arab Gulf and to the middle east.

English Page with different text

English Page with same text

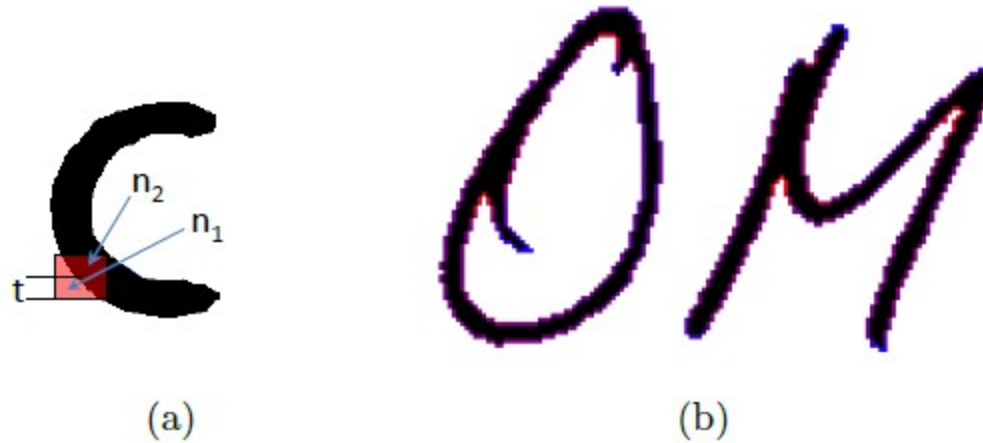
# Features-7066/doc

## *Directions*



**Fig. 1.** Computing local directions. (a) The predefined order for traversing shapes. (b) Example of an ordered shape. (c) Estimating directions by computing linear regression of neighboring pixels. (d) Binary image and its corresponding Zhang skeleton: the red color corresponds to a  $\frac{\pi}{2}$  tangent and the blue color corresponds to a zero tangent.

# Curvatures



**Fig. 2.** (a) Computing curvatures. (b) Curvatures highlighted on binary image: red corresponds to the maximum curvature and blue corresponds to the minimum one.

## *Tortuosity*



**Fig. 3.** Computing tortuosity (a) Longest traversing segment for 4 different pixels. (b) Length of maximum traversing segment: red corresponds to the maximum length, blue to the minimum one.

Chain code features, edge based directional features, ...

# Implementation

- **Number of Features/Dimensions:** 7066
- **Training Data:** 200 people
- **Testing Data:** 82 people

## Training Results:

- Using LSSVM on raw data – 78% correct predictions
- Using 2 Layer feed forward network with 20 hidden neurons – 80% correct predictions.  
0.54 log loss on Kaggle

# Problems with results:

- Because of too many dimensions, but limited number of samples, our training methods were very prone to Over-Fitting.
- So, we went for dimension reduction and tried the following techniques:
- PCA, Fisher Reduction, Kruskal Wallis Analysis and Minimum Redundancy Maximum Relevance(MRMR) Feature selection.

# Feature Reduction

This is how we fared:

- PCA: No improvement
- Fisher & Kruskal Wallis Analysis: Emphasis on *Curvature and Direction related features*, with moderate emphasis on Direction. No improve in accuracy by either SVMs or Neural Network method.
- MRMR: Strong Emphasis on *Chaincode and Curvature related features* , with moderate emphasis on Direction. Accuracy by SVMs and Neural Network Methods: 82%



# Observations:

- Chaincode metrics tend to be higher for females and lower for males, same goes for Curvature.
- Direction metrics do not have a general variation trend but vary across individual parameters.

Thank You for Your Attention.