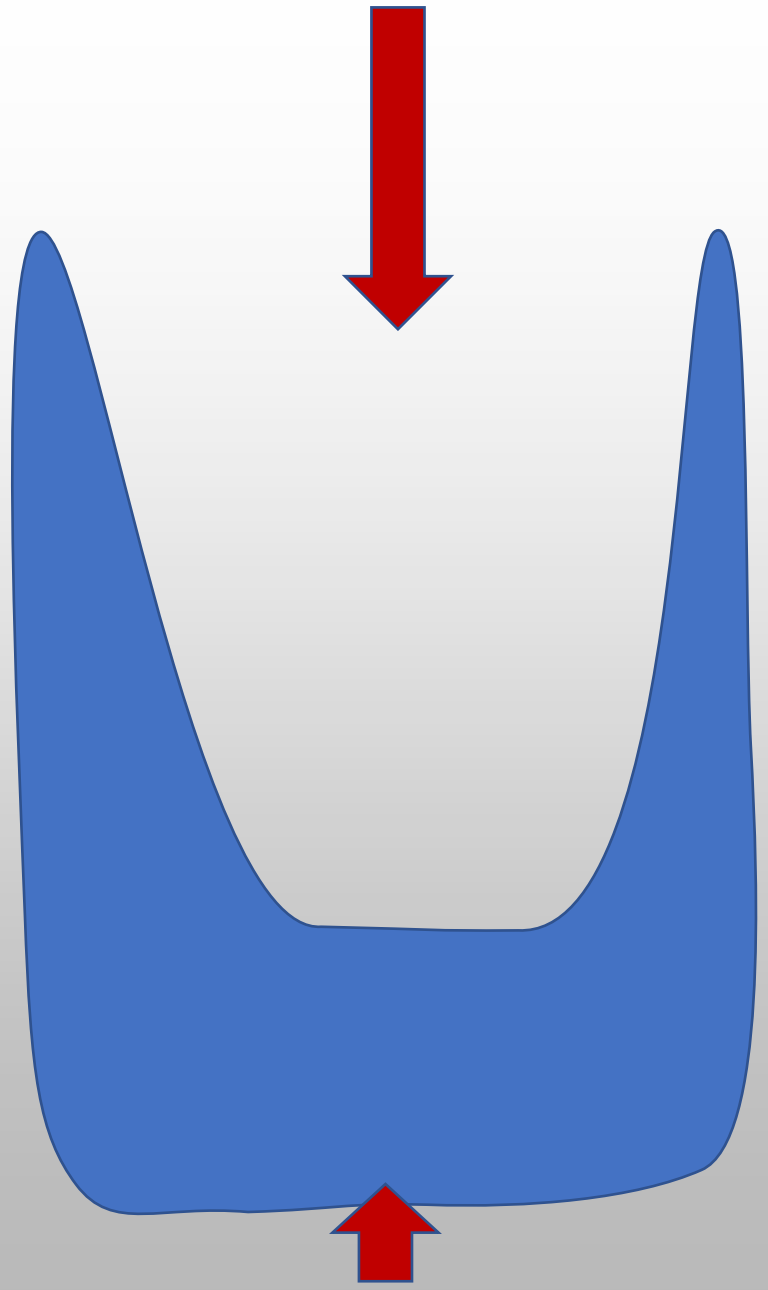
A karakterfelismerés buktatói

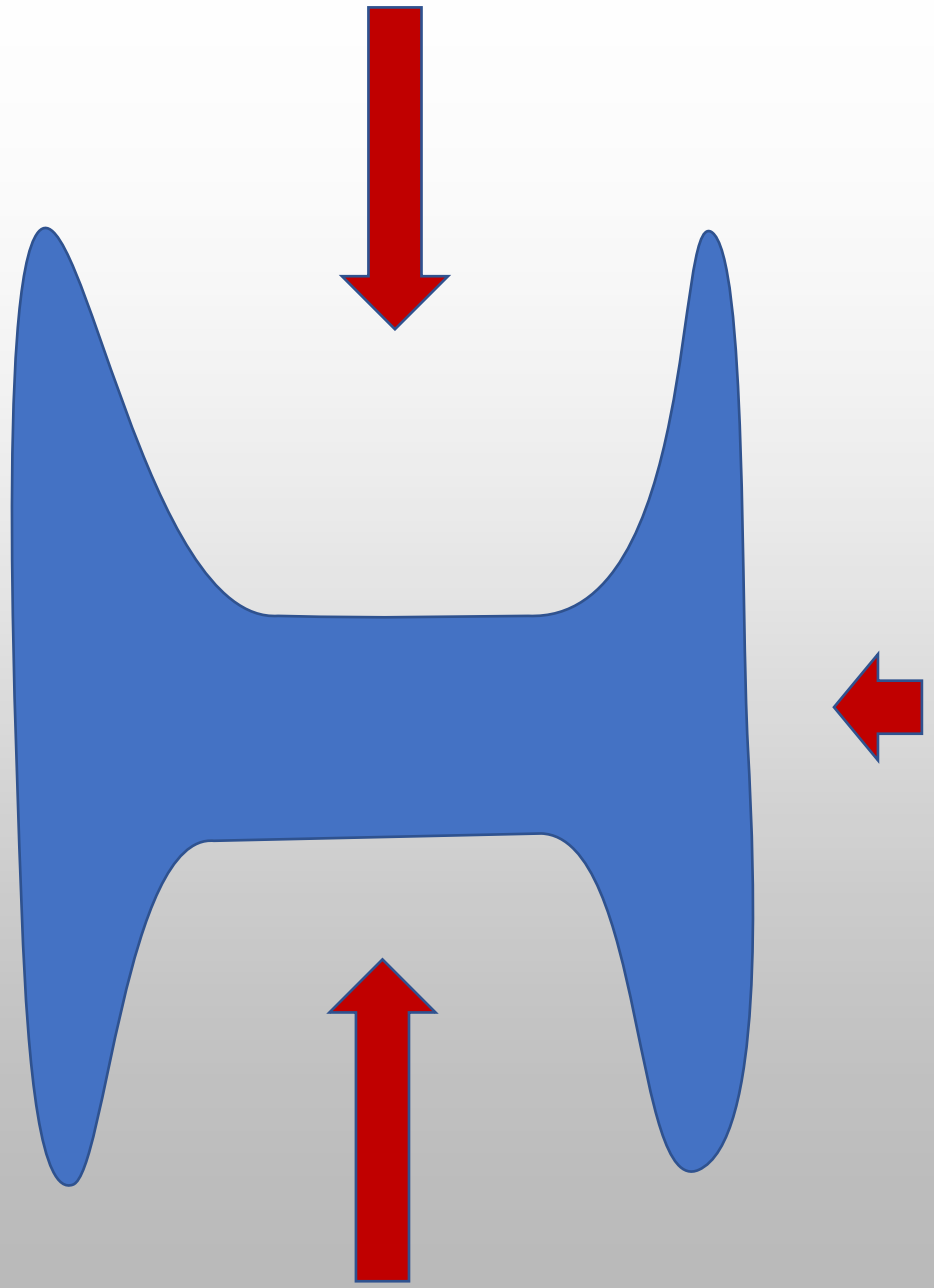
Szakmai szekció –
Mit főzünk a fazékban?

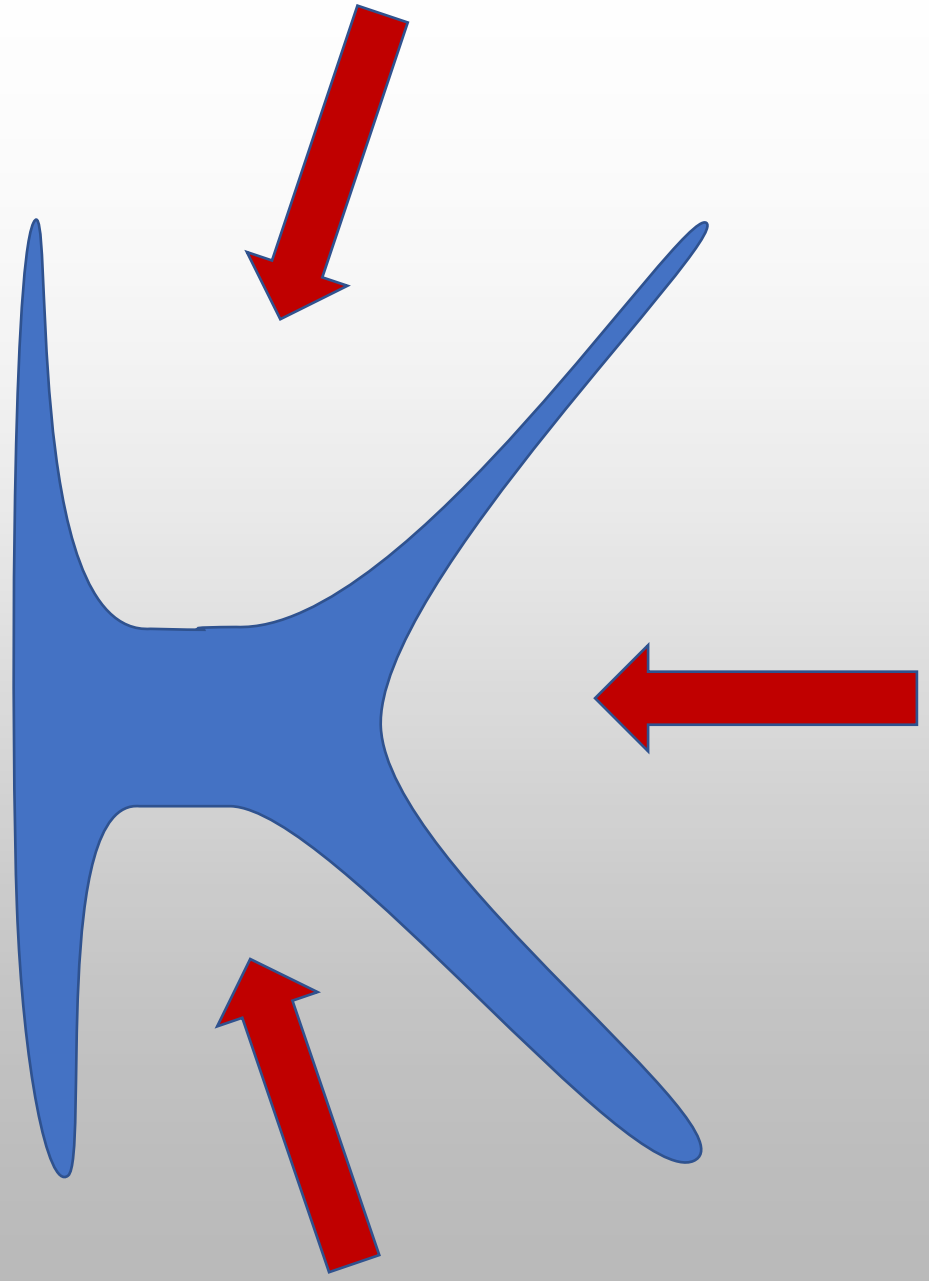
dr. Marosi István

A karakterfelismerés buktatói

- A kezdetek – 35 éve
  - 1984, Kovács Emőke, SzKI
    - M08X számítógép
    - Kamera felbontás: 384×288 pixel
    - Megvilágítástól erősen függ a kép
    - Nincs éles kép
    - Matrix-matching nem járható út
  - Ötlet: gyurma

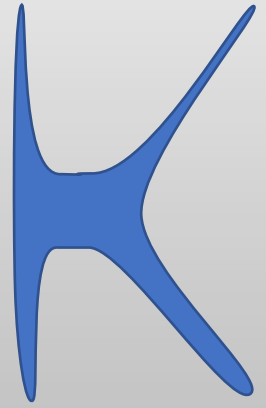


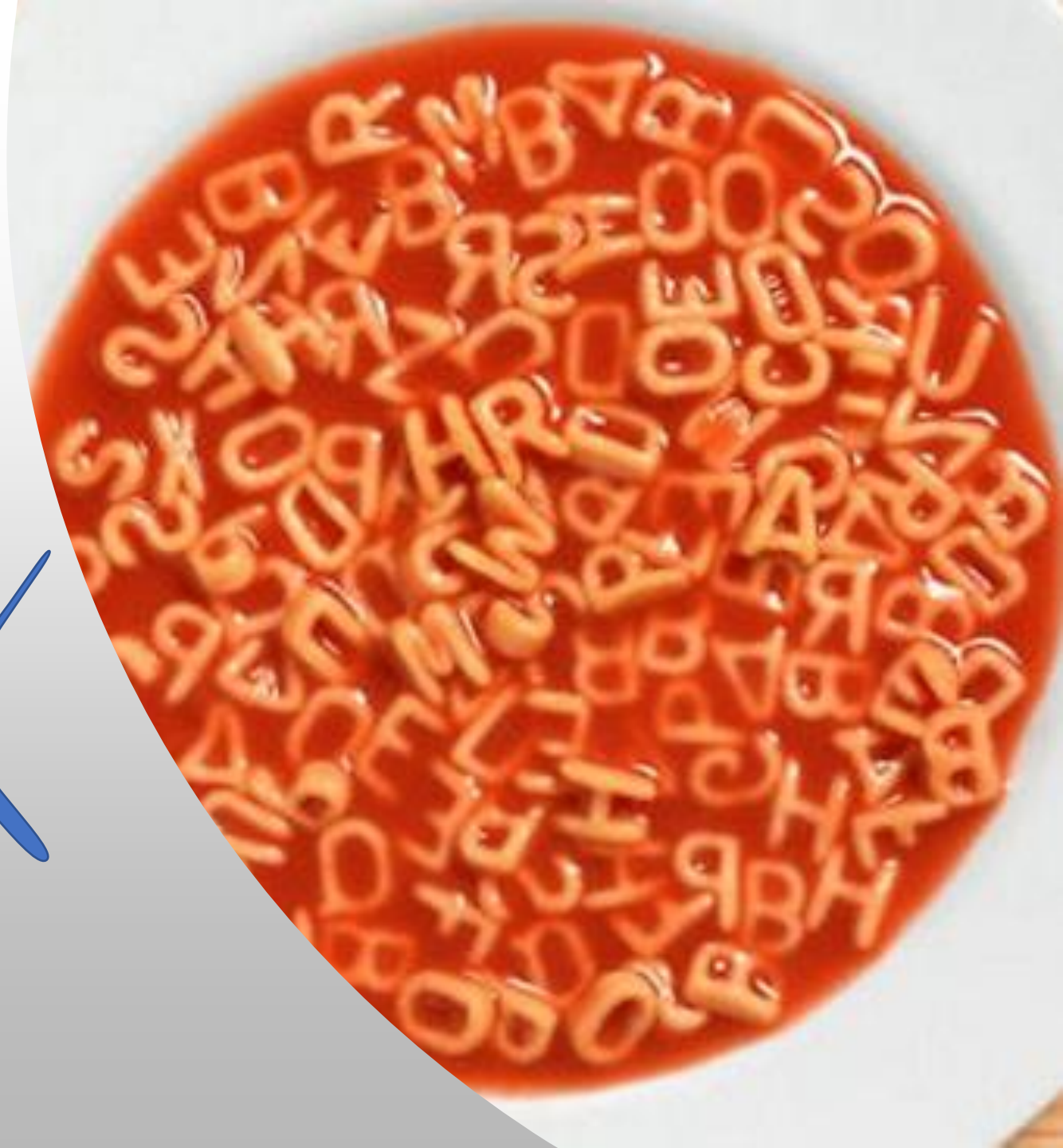




A karakterfelismerés buktatói

- Feature-ök:
  - A konkáv öblök helye és iránya
- Kézzel rajzolt karakter-kártyák
- Első program:
  - Kontúrkövetés alapú
  - Bedrótozott logika
  - If blokkok sorozata





A karakterfelismerés buktatói

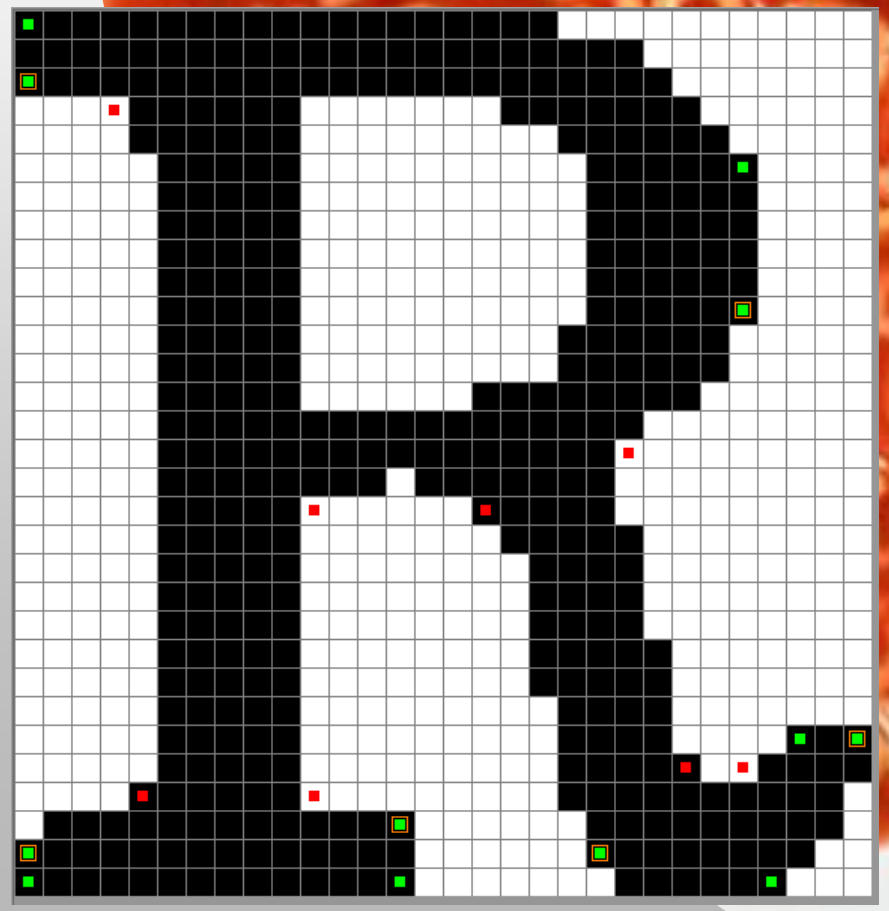
- 1985:
  - Kovács Emőke, Marosi István
- Új hardware
  - Microtek scanner
    - MS-300A: 300 dpi BW
  - IBM PC klón (Proper-16)
  - Multitech grafikus kártya
    - 1024×768 monochrome
    - (CGA BW: 640×200)
- Grafikus – ablakos interfész
  - BitBlt műveletek
  - Egér nincs – szimulátor
  - Scanner és BMP file kezelés
  - ASM-ben írtuk
    - 1985-ben természetesnek tűnt

A karakterfelismerés buktatói

- 16 bites IBM PC
  - Csak 64k-t lehet könnyen címezni
  - Egy A4-es BW oldal 1.04 MB
    - Csak egy része fér be!
    - 128 képsor, gyűrűpuffer
- 1986: Első „nagy” bug:
  - Program > 32k
- 1996: 32 bites kód
  - movlp es::di,cx::dx
    → mov edi,edx
- 2012: 64 bites kód
  - Ugyanabból a forrásból 3 architektúra
    → mov rdi,rdx

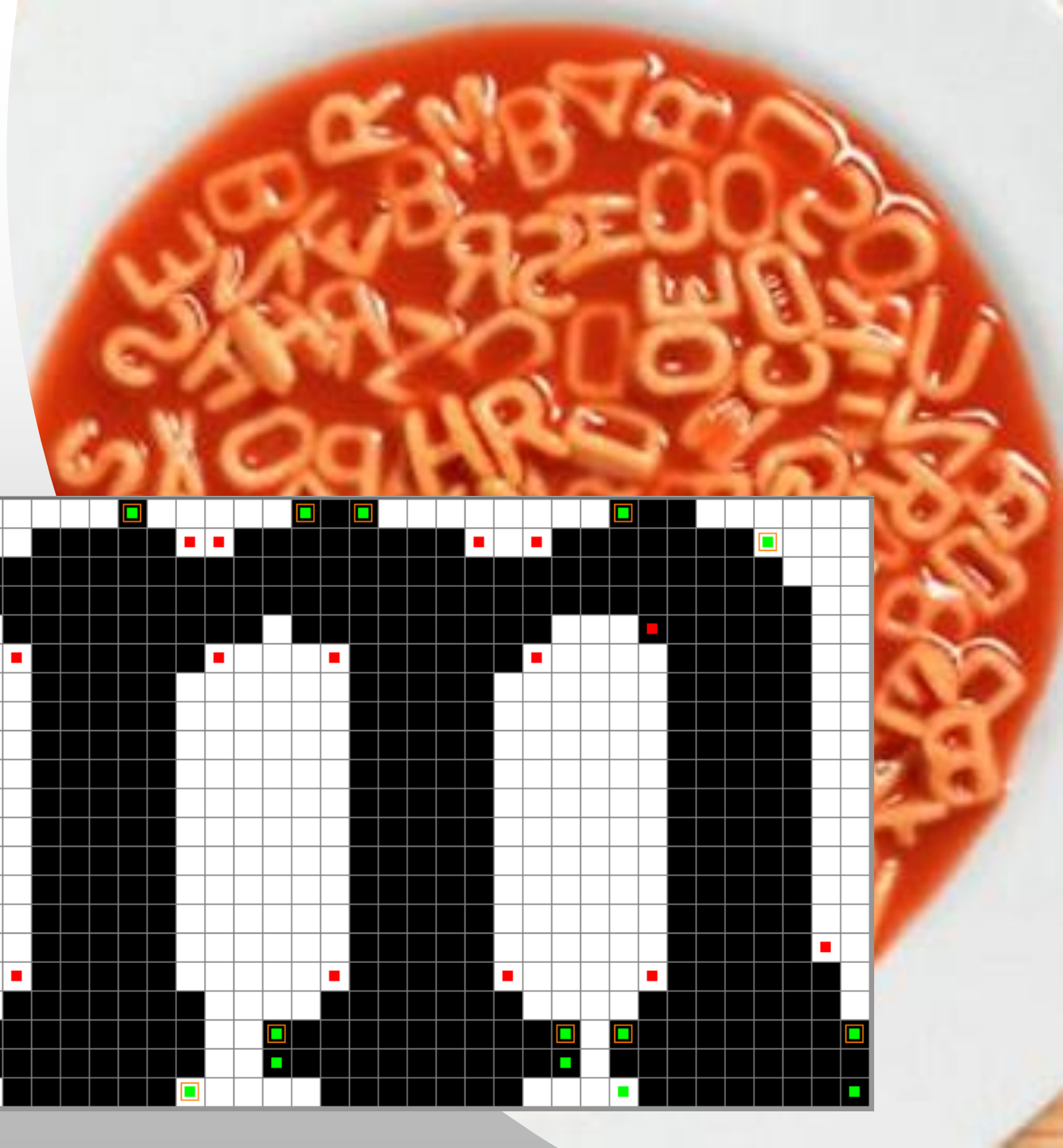
A karakterfelismerés buktatói

- Feature-ök (1985):
  - Konvex (zöld) és konkáv (piros) töréspontok helye a kontúron
    - A pontok száma az elsődleges feature!
- Betanított betűalakok
  - Nincs bedrótozva semmi
  - K-Nearest Neighbors osztályozó
  - Egyszerű tanítás
    - PAT file: címkézett pattern gyűjtemény
    - Tanuló program összevonja a hasonlókat
  - Viszonylag kevés alak is elegendő



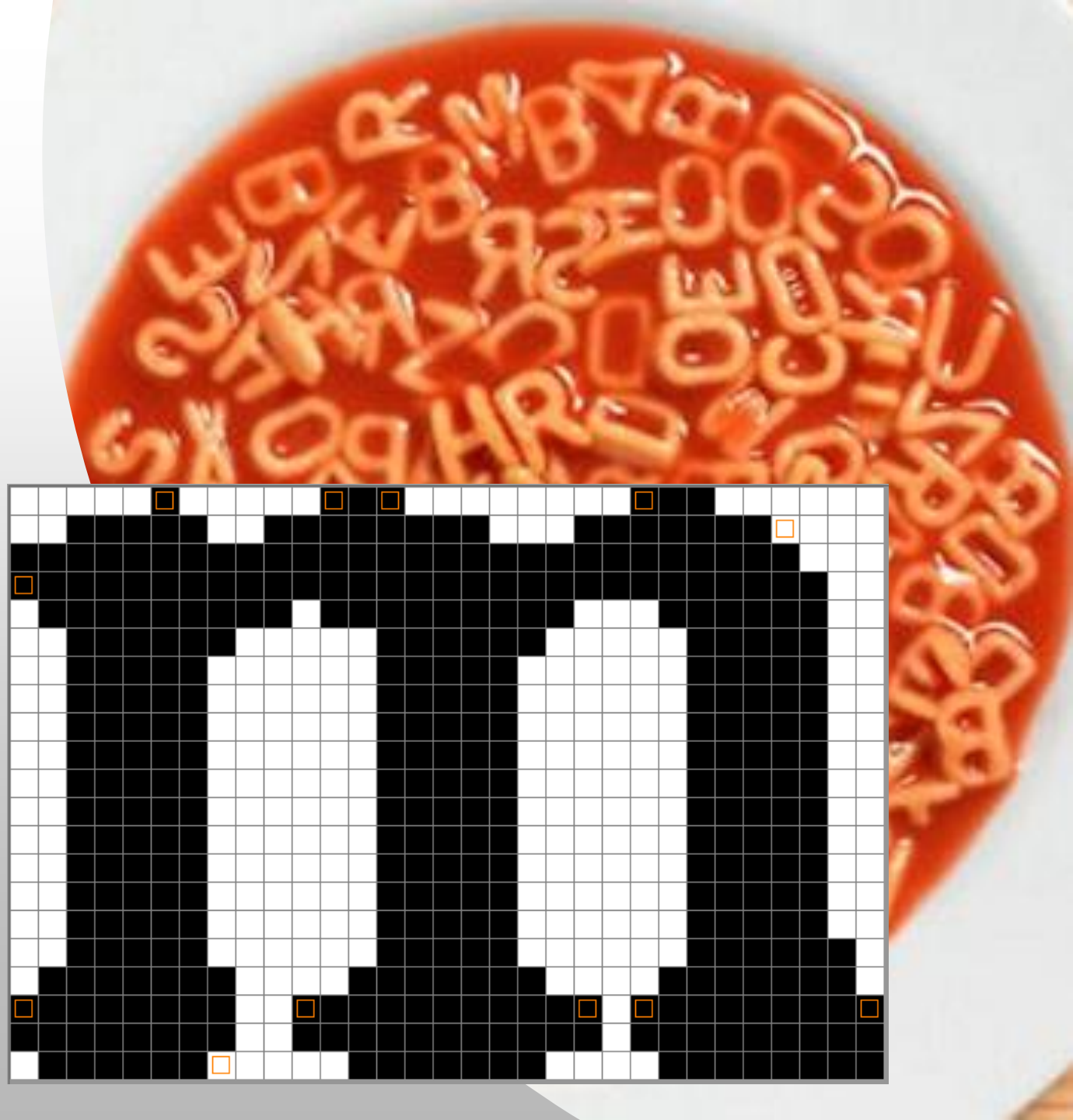
A karakterfelismerés buktatói

- Mennyi a „viszonylag kevés” alak?
  - Néha meglepően sok:
    - Talpak hajlamosak eltűnni
    - Mindig kimarad valamilyen alak
      - Érzékeny pontok, 2^n alak generálás
- Feature-ök 2.0:
  - Töréspontok helyett ívek



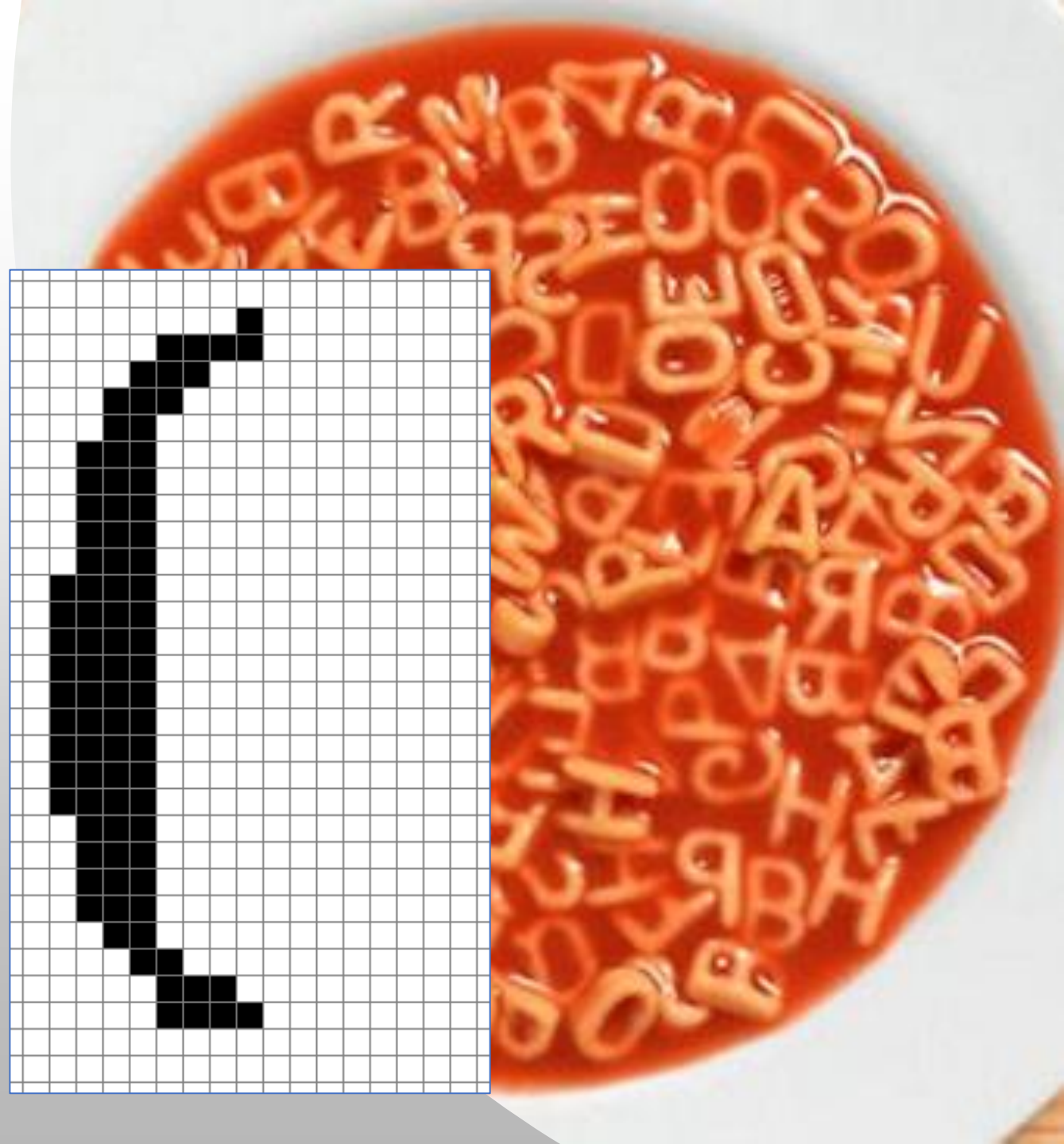
A karakterfelismerés buktatói

- Mennyi a „viszonylag kevés” alak?
  - Néha meglepően sok:
    - Talpak hajlamosak eltűnni
    - Mindig kimarad valamilyen alak
      - Érzékeny pontok, 2^n alak generálás
- Feature-ök 2.0 (1990):
  - Töréspontok helyett ívek
    - Sokkal kevesebb változatosság
    - Viszont gyengébb alak-leírás
      - ISIT rutinok szükségesek



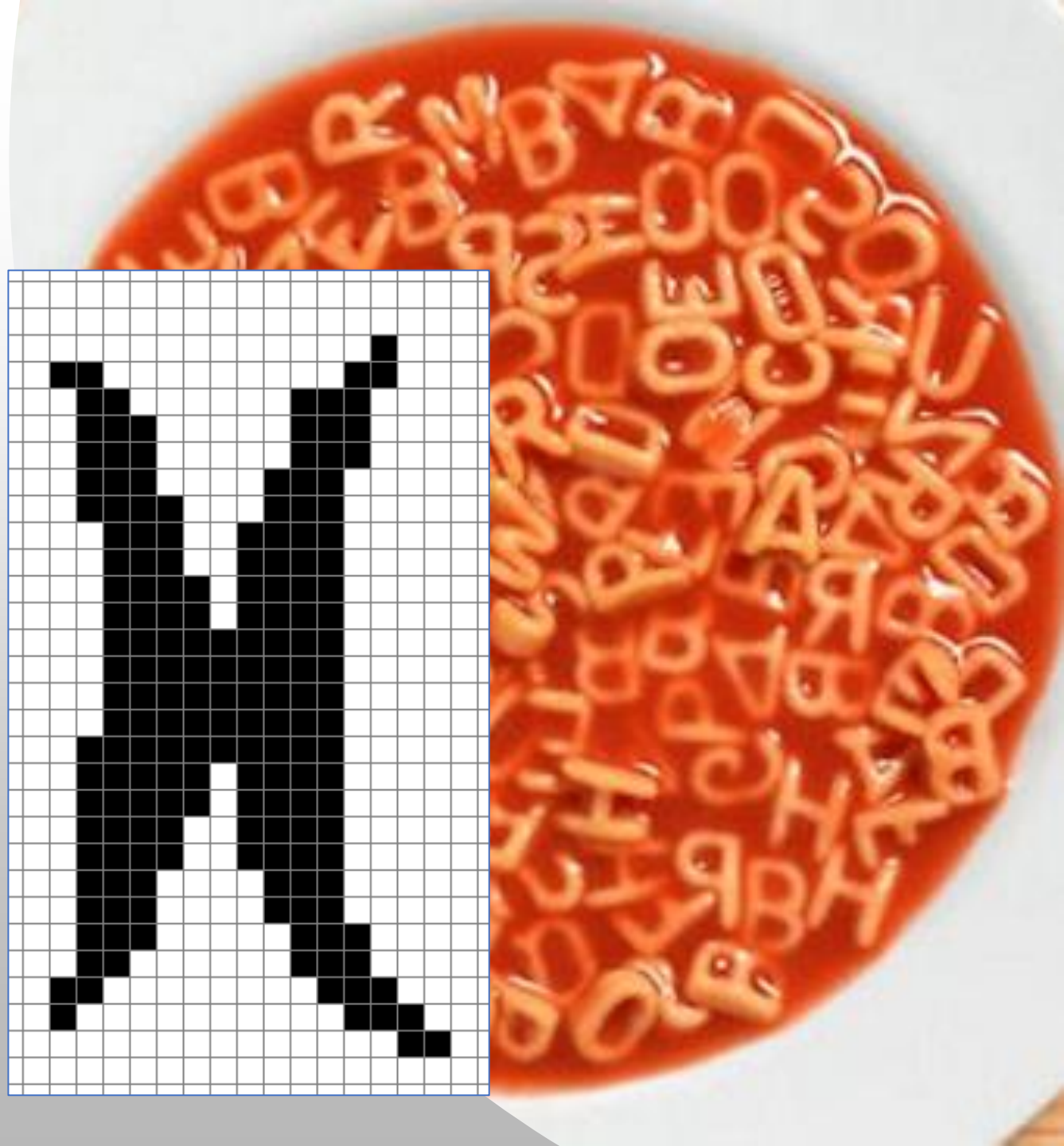
A karakterfelismerés buktatói

- Nagyon egyszerű alakok
  - Más feature is kell hozzájuk



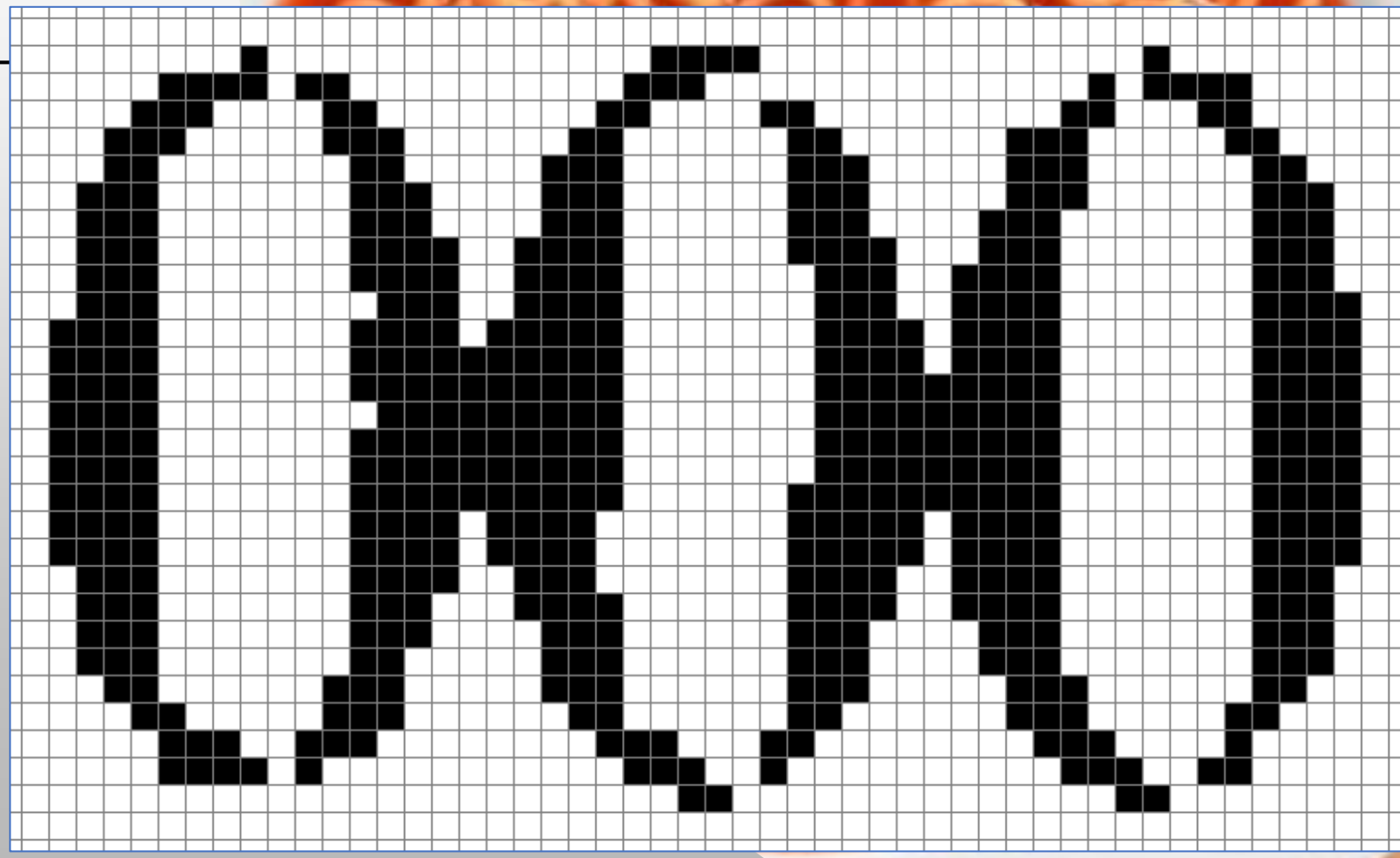
A karakterfelismerés buktatói

- Nagyon egyszerű alakok
  - Más feature is kell hozzájuk
- A program így látja, pixel-közelből



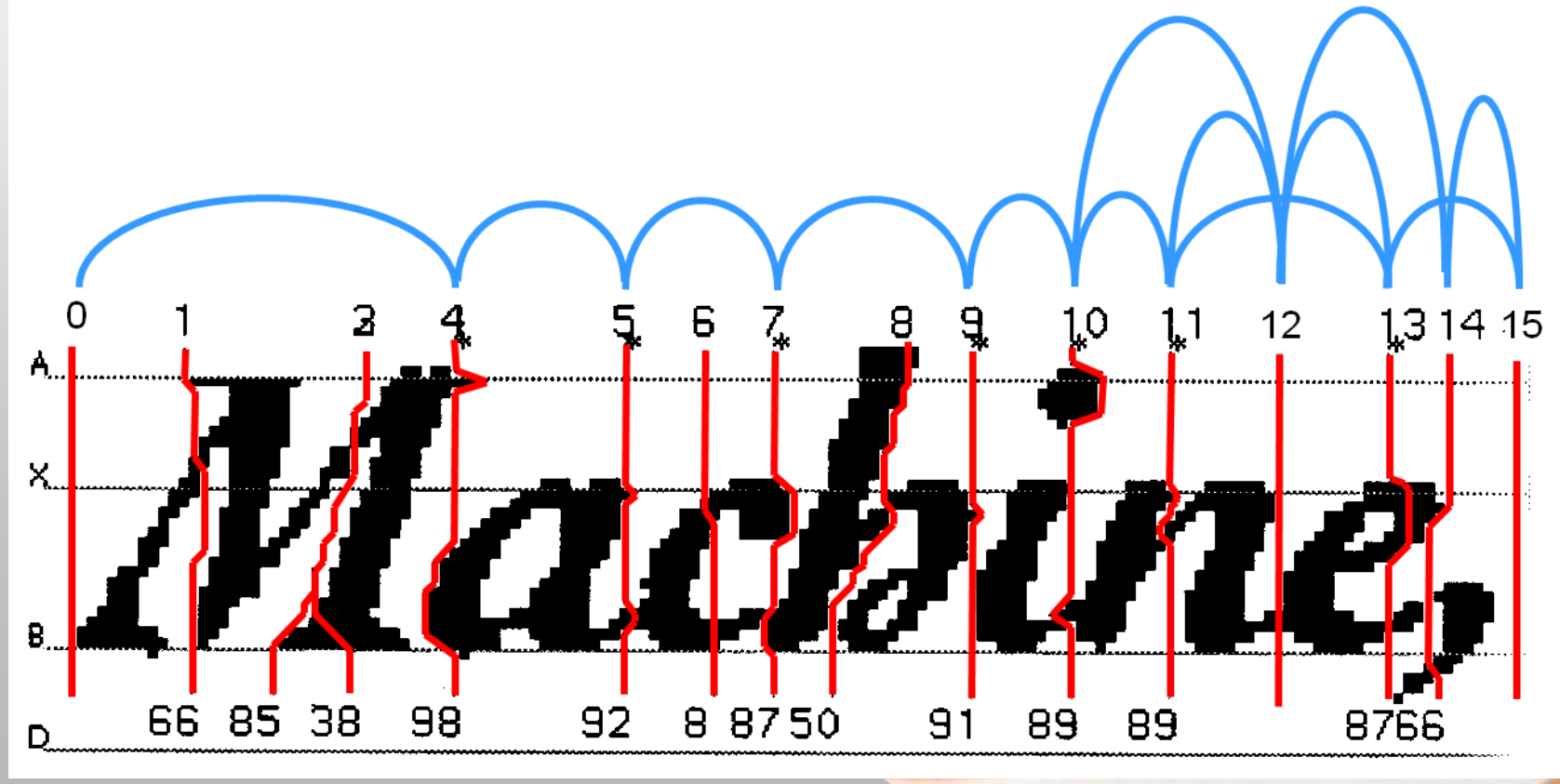
A karakterfelismerés buktatói

- Karakter szegmentálás
  - A felismerési problémák nagy része szegmentálás alapú!



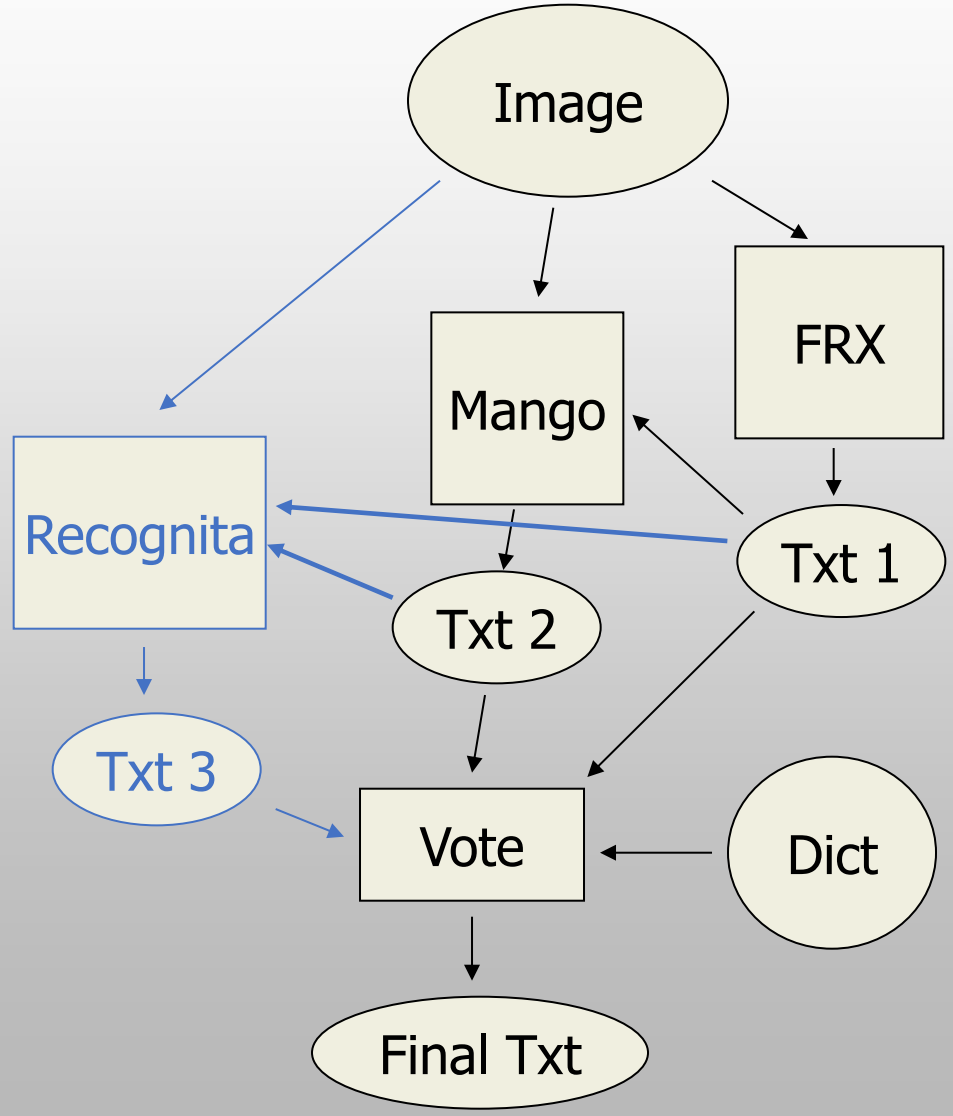
A karakterfelismerés buktatói

- Karakter szegmentálás
  - A felismerési problémák nagy része szegmentálás alapú!
  - Iteratív processz, sok lehetséges szegmentálást kell kipróbálni, felismerni



Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
  - 5 OCR engine:
    - 4 nyugati: Recognita, M-Text, Fireworx, Mango
      - Mindegyikben:
        - Elő-tanított osztályozás
        - Szegmentálás
        - Adaptív osztályozás
        - Adaptív szegmentálás
        - Nyelvi analízis
      - Recognita és Mango:
        - Belső vote-olás az algoritmusok mélyén
    - 1 ázsiai: kínai partnertől



Az OCR nem csak karakterfelismerés!

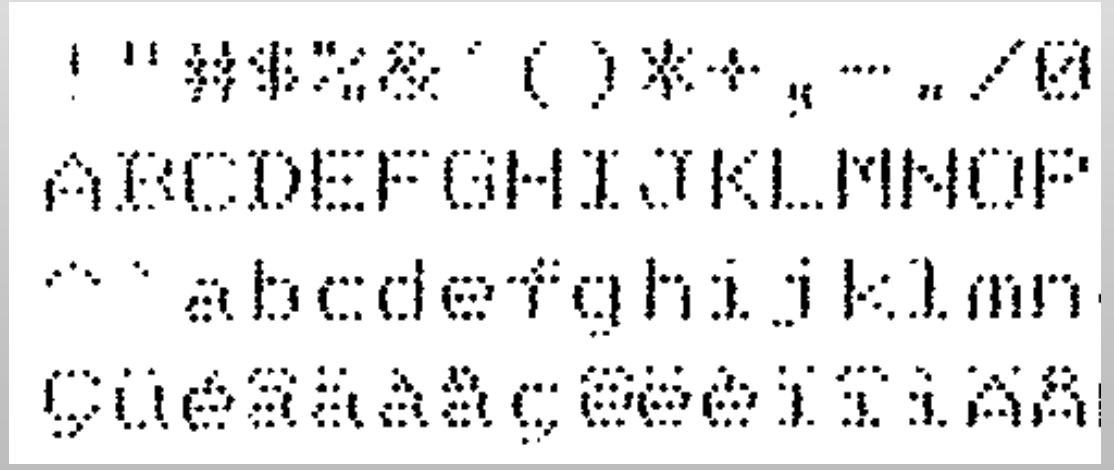
- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
  - 5 OCR engine
  - Kép beolvasás:
    - Scanner kezelés
    - Képfile-ok
    - PDF file-ok

Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
  - 5 OCR engine
  - Kép beolvasás
  - Kép előfeldolgozás:
    - 2D/3D de-skew, de-speckle, mindenfajta szűrés és bűvészkedés
    - Sorkiegyenesítés
    - Binarizálás

Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
  - 5 OCR engine
  - Kép beolvasás
  - Kép előfeldolgozás
  - Mindenféle felismerés:
    - Barcode
    - OMR
    - Pontmátrix
    - Kézzel írt számok
    - Nyelv felismerés

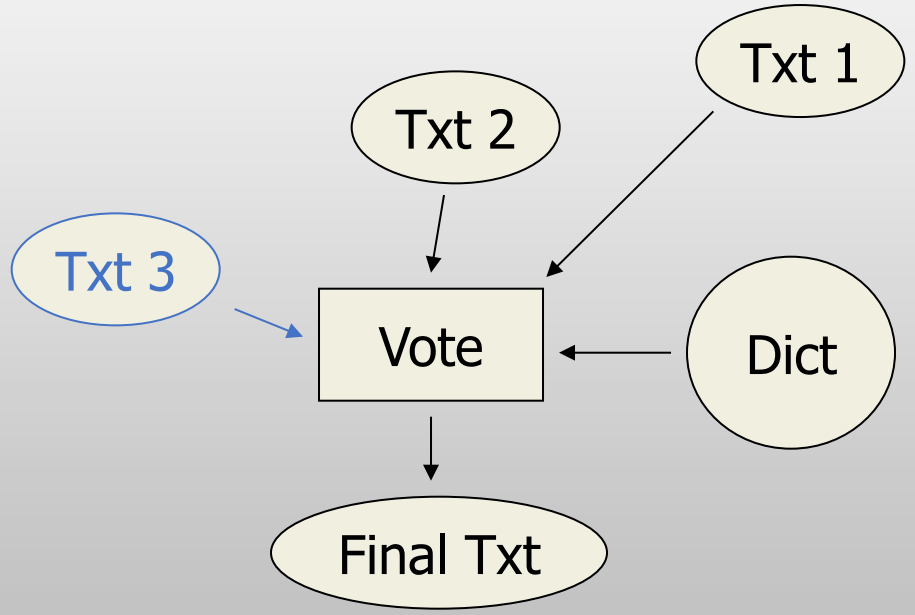


Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
  - 5 OCR engine
  - Kép beolvasás
  - Kép előfeldolgozás
  - Mindenféle felismerés
  - Zónázás:
    - Hol vannak a szövegek, táblázatok, grafikák?

Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
  - 5 OCR engine
  - Kép beolvasás
  - Kép előfeldolgozás
  - Mindenféle felismerés
  - Zónázás
  - Vote-olás:
    - Külső vote-olás:
      Az engine-ek outputja alapján egy jobb eredmény kialakítása



Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
  - 5 OCR engine
  - Kép beolvasás
  - Kép előfeldolgozás
  - Mindenféle felismerés
  - Zónázás
  - Vote-olás
  - Form-ok kezelése:
    - Form elemek felismerése
      - Kitöltendő mező
      - Check mark
      - „Fésű”
    - Template illesztés

PTX RESEARCH

PTX is an affirmative action/equal employment opportunity employer. Discrimination because of race, color, religion, sex, handicap, sexual orientation or national origin is prohibited. In order to be considered for an internship, you must submit a completed application form along with a cover letter and your resume.

Name(s) of Internship(s) Applied For:

1. Internship1 3.
2. Internship2 4.

Contact Information

Name: Laszlo Fero

Address: Fay Andras u 35/A

City: Pomaz State: Zip Code: 2013

Telephone Numbers: (Home) +3628123466 (Mobile) +3620346879

Email Address: laszlo.fero@nuance.com

Are you legally eligible to work in the U.S.? ☑ Yes ☐ No

Are you requesting college credit hours for your internship? ☑ Yes ☐ No

If Yes, College Name:

If you do not receive an internship at PTX, would you be interested in a fee-based program?

☐ Yes ☑ No

Education

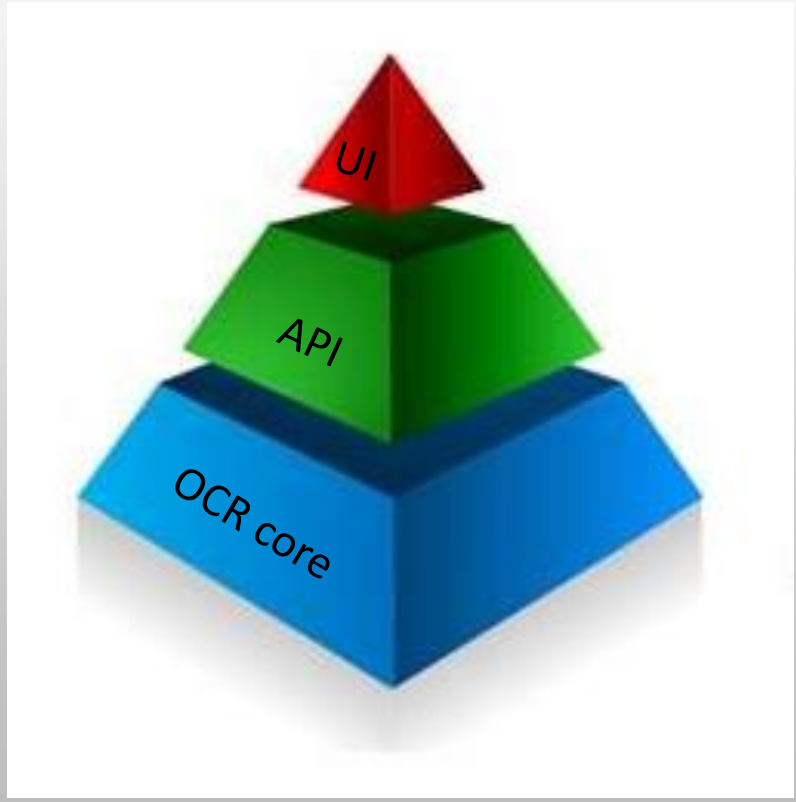
TYPE OF SCHOOL	NAME AND LOCATION	DEGREE/DATE	MAJOR
High School			
College			

Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
  - 5 OCR engine
  - Kép beolvasás
  - Kép előfeldolgozás
  - Mindenféle felismerés
  - Zónázás
  - Vote-olás
  - Form-ok kezelése
  - Output dokumentum előállítás:
    - PDF processzálás
    - Több lapos formattált dokumentumok (pl. docx, xlsx, html)

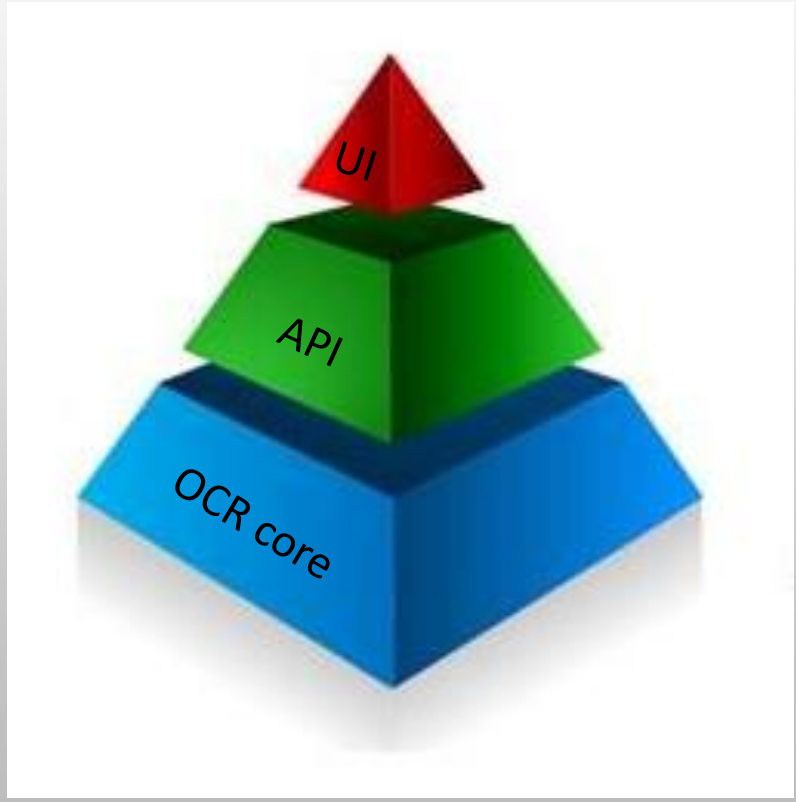
Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
  - 5 OCR engine
  - Kép beolvasás
  - Kép előfeldolgozás
  - Mindenféle felismerés
  - Zónázás
  - Vote-olás
  - Form-ok kezelése
  - Output dokumentum előállítás
  - Interfészek:
    - Windows UI
    - C, C#, Java API



Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
  - 5 OCR engine
  - Kép beolvasás
  - Kép előfeldolgozás
  - Mindenféle felismerés
  - Zónázás
  - Vote-olás
  - Form-ok kezelése
  - Output dokumentum előállítás
  - Interfészek
  - És még sok más…



Köszönöm a figyelmet