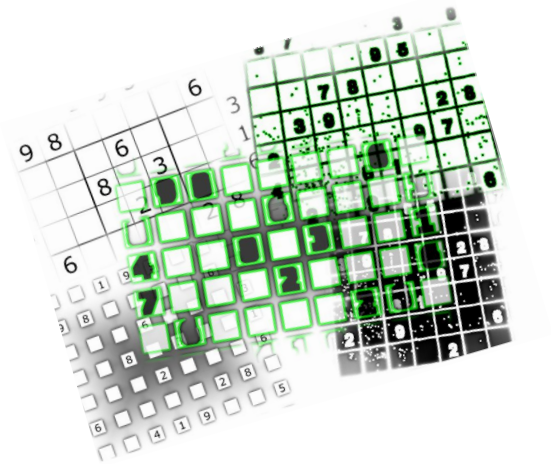


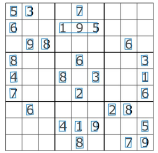
~~Sudoku~~ ~~and~~ Optical Character Recognition ~~in~~ ~~Python~~

IDA Østjylland Seminar, 23/10/2023

Luminita C. Totu
Control Engineer
luminita.totu@gmail.com

ChatGPT
OpenAI standard interface
chat.openai.com

Wikipedia
https://en.wikipedia.org/



Optical **C**haracter **R**ecognition - What is It ?

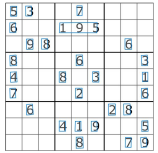
- An advanced image analysis process where patterns of pixels, corresponding to letters, numbers, or other symbols within a digitized image, are systematically identified and classified,

- Image can be: scanned/photographed various printed material (documents from financial, medical, passports, books, newspaper etc), handwrite, or other “wild” text (road signs, building signs, number plates) aka **S**cene **T**ext **R**ecognition)

- It is a problem a bit older than computers

IBM 1418 in Postgirots datacenter

“IBM designed, manufactured and sold optical mark and character readers from 1960 until 1984. The IBM 1287 is notable as being the first commercially sold scanner capable of reading handwritten numbers.”



Optical **C**haracter **R**ecognition - What is It ?

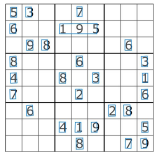
“Early optical character recognition may be traced to technologies involving telegraphy and creating reading devices for the blind”

- 1914, Emanuel Goldberg - developed a machine that read characters and converted them into standard telegraph code (?)

- 1920, Emanuel Goldberg - "**Statistical Machine**" for searching microfilm archives using an optical code recognition system (?)

- 1913, Edmund Fournier d'Albe **Optophone** - a handheld scanner that when moved across a printed page, produced tones that corresponded to specific letters or characters (search also “Stereotoner” on YouTube)

- 1970, John Linvil, **Optacon** (tactile) - various materials on YouTube



Optical **C**haracter **R**ecognition - What is It ?

- Farrington 7B font numeric font used on credit cards (David H. Shepard)

- OCR-A and OCR-B are typefaces (or fonts) that were specifically created for Optical Character Recognition (OCR) systems. Both these fonts were designed to be easily read by both machines and humans, which was essential in the early days of OCR

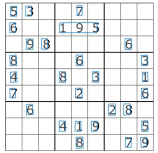
The Quick Brown
Fox Jumps Over
The Lazy Dog

abcdefghijklmnopqrtuvwxyz0123456789 [ ] ( ) { } / \ < >

The Quick Brown
Fox Jumps Over
The Lazy Dog

abcdefghijklmnopqrstuvwxyz0123456789 [ ] ( ) { } / \ < >

OCR-A (left) and OCR-B (right) samples wikipedia)



Optical **C**haracter **R**ecognition - What is It ?

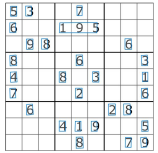
The evolution of OCR reflects the broader trajectory of technology: from narrow, specific tasks in controlled environments to versatile, real-time applications in diverse, real-world scenarios.

Early-days:

- Basic Recognition - focus was on recognizing standard printed fonts
- Standardized Documents - early applications targeted specific, standardized documents like bank checks and telegrams.
- Manual Correction - a necessary step

Mid-Phase Evolution:

- Expanded Font Recognition
- Handwriting Recognition
- Multi-language Support
- Improved Accuracy
- Commercial OCR software packages

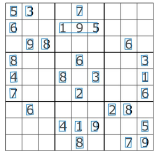


Optical **C**haracter **R**ecognition - What is It ?

The evolution of OCR reflects the broader trajectory of technology: from narrow, specific tasks in controlled environments to versatile, real-time applications in diverse, real-world scenarios.

Present-Day Challenges and Focus:

- **Advanced handwriting recognition**
- **Real-Time processing** - with mobile devices, there's a push for real-time OCR, e.g., reading restaurant menus or signs through augmented reality apps.
- **Complex Documents** - modern OCR systems aim to interpret not just text, but also the structure and layout of complex documents, including tables, graphics, and multi-column formats.
- **Low quality scans** - OCR systems now aim to decipher text from low-resolution or degraded documents, such as old newspapers or worn-out books
- **Integration with AI** - Modern OCR is increasingly integrated with broader artificial intelligence systems for tasks like automatic data entry, information extraction, and context-aware processing
- **Inclusive Designs**: Recognizing text from diverse sources, ensuring that OCR works effectively across different cultural, linguistic, and design contexts.
- **Training Data:** Modern machine learning-based OCR systems require vast amounts of training data, leading to challenges in data acquisition and curation.



Optical **C**haracter **R**ecognition - State of the Art

- [Microsoft] Azure OCR Engine:

Easy to use online demo tool (**no account needed**):
https://portal.vision.cognitive.azure.com/demo/extract-text-from-images

- Google Vision AI

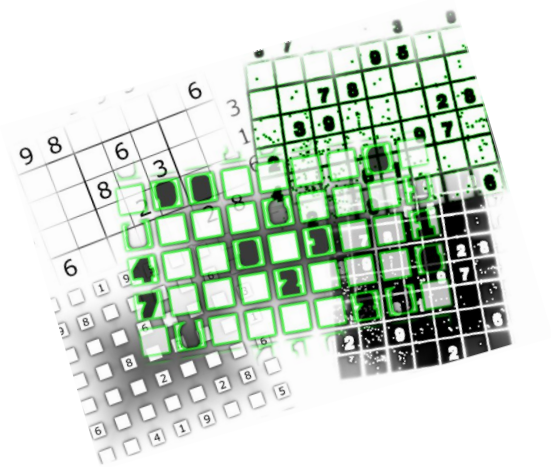
Requires Google Cloud account / Trial can be used but CC is needed
https://cloud.google.com/vision?hl=en (scroll to Demo “Try the API”)

- Amazon Textract

Requires AWS account / Trial can be used but CC is needed
https://aws.amazon.com/console/

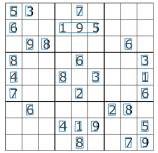
- OCR.space

Easy to use online demo tool **(no account needed)**
https://ocr.space/



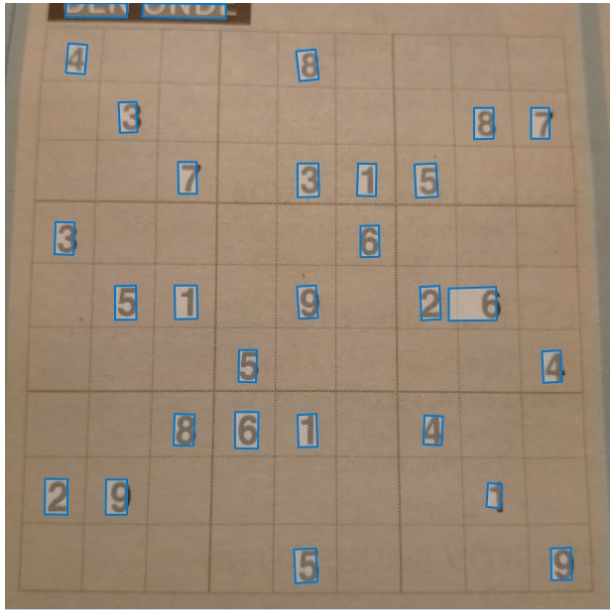
Sudoku and Optical Character Recognition ~~in Python~~

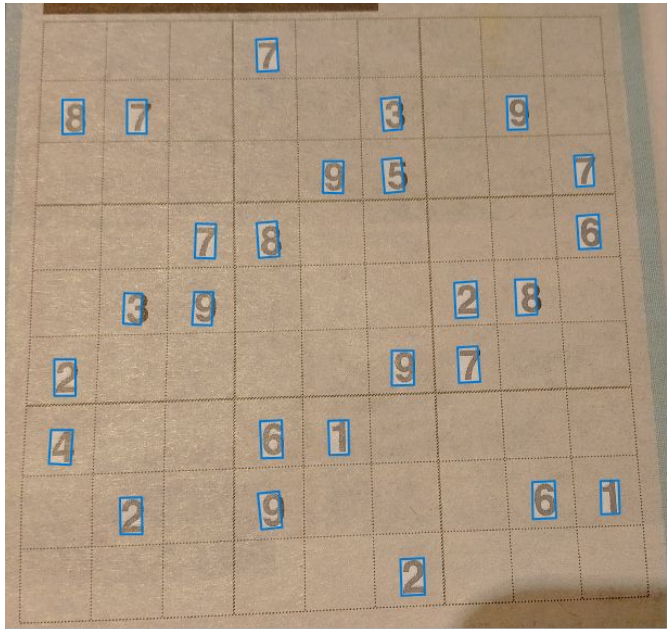
Activity !

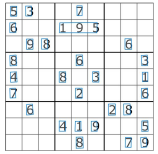


Optical **C**haracter **R**ecognition of Sudoku Grids

- [Microsoft] Azure OCR Engine ★★★★★

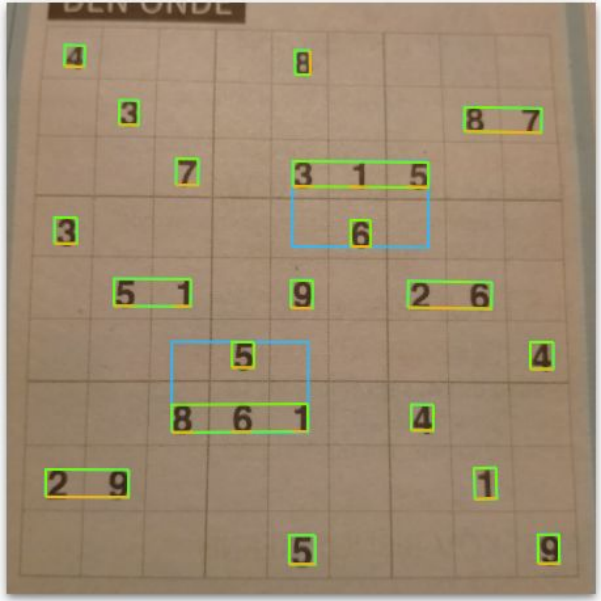


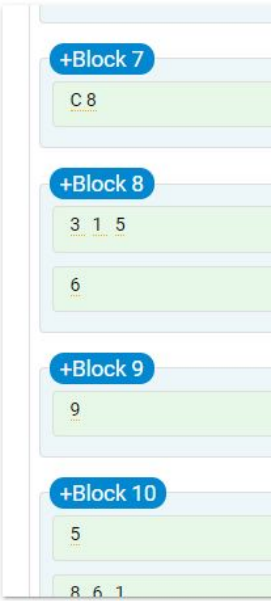




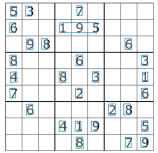
Optical **C**haracter **R**ecognition of Sudoku Grids

- Google Vision AI ★★★★



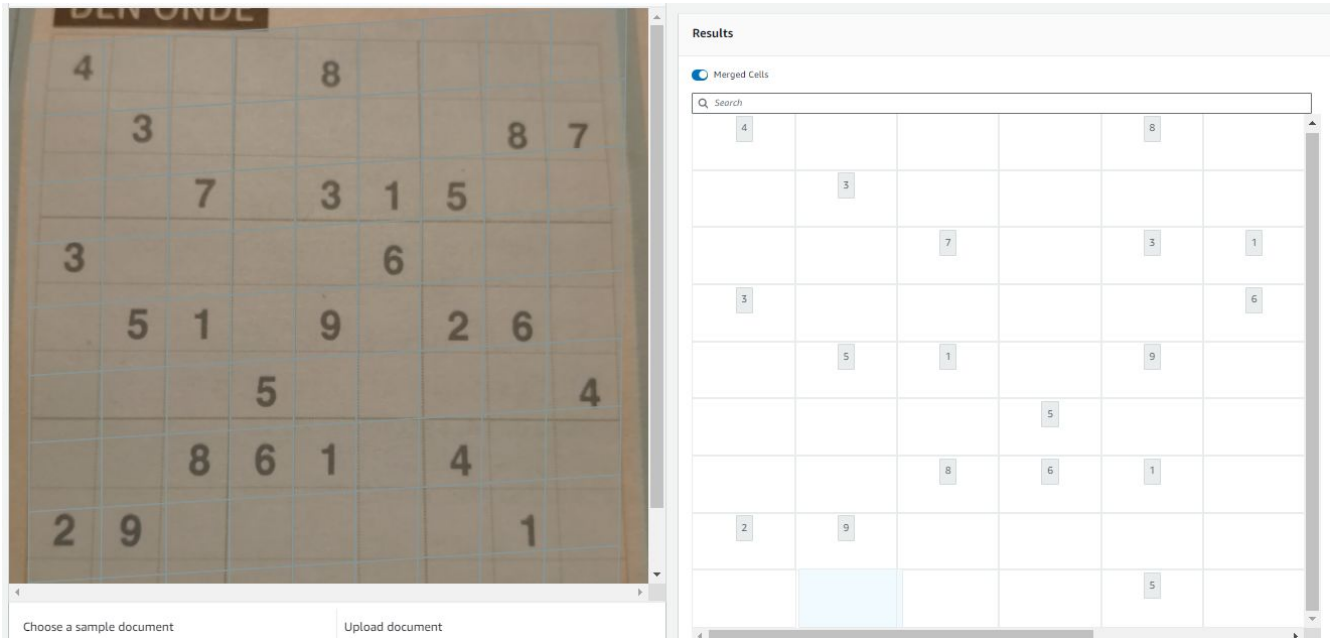


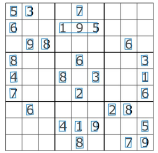
OCR mistake !



Optical **C**haracter **R**ecognition of Sudoku Grids

- Amazon Textract (table) ★★★★★





Optical **C**haracter **R**ecognition of Sudoku Grids

- OCR.space ★★★★ // Create searchable PDF with visible text layer (Engine 3)