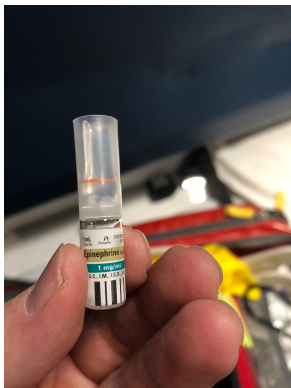


Optical Character Recognition for Pill Bottles

Ameera Chowdhury and NGFR Team at JPL

Jet Propulsion Laboratory

Accurate and Fast Communication Between EMTs



- In a medical emergency, it is crucial to correctly record a patient's medication and quickly communicate it to everyone on the medical team.
- Information about a patient's medication is often found on pill bottles at the scene.

Getting Information from Medicine Bottles

The diagram shows a sample medicine label for 'Local Pharmacy' with various fields and callouts. The label includes the pharmacy's name and address, a phone number, a doctor's name, a prescription number, a patient's name and address, instructions for use, the drug name and strength, and an expiration date. Callouts point to these fields with labels: 'Pharmacy name and address', 'Number used by the drugstore to identify this drug for your refills', 'Person who gets this drug', 'Instructions about how often and when to take this drug', 'Name of drug and strength of drug', 'Number of refills before certain date', 'Doctor's name', 'Drugstore phone number', 'Prescription fill date', and 'Don't use this drug past this date'.

Pharmacy name and address

Number used by the drugstore to identify this drug for your refills

Person who gets this drug

Instructions about how often and when to take this drug

Name of drug and strength of drug

Number of refills before certain date

Doctor's name

Drugstore phone number

Prescription fill date

Local Pharmacy
123 MAIN STREET
ANYTOWN, USA 11111
(800) 555-5555

DR C. JONES

NO 0060023-08291 DATE 06/23/09

JANE SMITH
456 MAIN STREET ANYTOWN, US 11111

TAKE ONE CAPSULE BY MOUTH THREE TIMES DAILY FOR 10 DAYS UNTIL ALL TAKEN

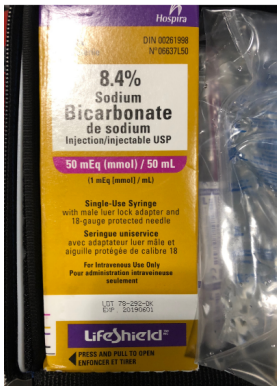
AMOXICILLIN 500MG CAPSULES

QTY MFG
NO REFILLS - DR. AUTHORIZATION REQUIRED
USE BEFORE 06/23/12
SLF/SLF

Don't use this drug past this date

Goal: From a photo of a medicine bottle, glean important information such as the name of the medication, the dose, the lot number, and the expiration date.

Real Example



Hospira

DIN 00261998

N° 06637 L50

8.4%

Sodium

de sodium

50 mEq (mmol)/50 mL

Injection/injectable USP

(1 mEq [mmol] /mL)

Single-Use Syringe

with male luer lock adapter and

18-gauge protected needle

Seringue uniservice

avec adaptateur luer mâle et

aiguille protégée de calibre 18

For Intravenous Use Only

Pour administration intraveineuse

seulement

LIT 8-292-0

EF 2019ne01

TM/

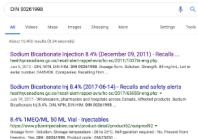
MC

PRESS AND PULL TO OPEN

ENFONCER ET TIRER

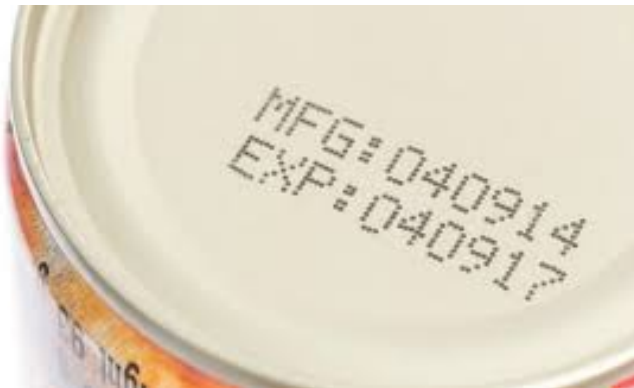
DIN 00261998 Sodium

Bicarbonate



- The DIN number can identify the medication.
- Lot numbers and expiration dates are difficult for OCR engines because they are written in a dot matrix font.

Challenges



- No publicly available dataset for dot matrix fonts exists.
- Need to generate a dot matrix font dataset from scratch.
- Must train an OCR engine, like Tesseract, ourselves.

Constructing the Dataset



Generating a large dot matrix font dataset by

- randomly shifting dots
- deleting dots
- rotating and shearing



Python, LaTeX, Keras ImageDataGenerator, Tesseract