

# Riconoscimento automatico di libri di letteratura fantastica

*di Giacomo Fantazzini*

[giacomo.fantazzini2@studio.unibo.it](mailto:giacomo.fantazzini2@studio.unibo.it)

# Il Problema:

Libri con stesso contenuto,  
ma titolo differente

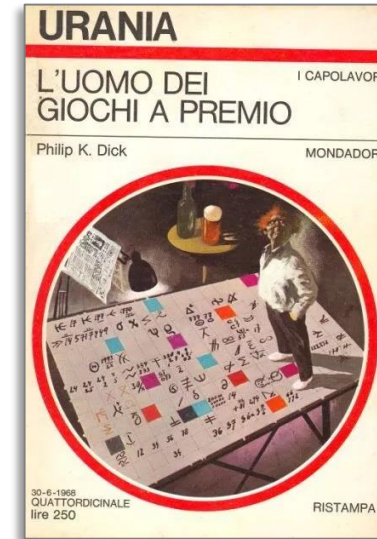
Libri con stesso titolo,  
ma contenuto differente

## Obiettivo:

Realizzare un'app che  
identifichi edizione e  
contenuto di un libro a  
partire da una foto della  
copertina



**Il tempo si è spezzato**  
( *Time Out of Joint* )  
Philip K. Dick  
1959



**L'uomo dei giochi  
a premio**  
( *Time Out of Joint* )  
Philip K. Dick  
1968  
1975  
1983



**Tempo fuori luogo**  
( *Time Out of Joint* )  
Philip K. Dick  
1996  
1999



**Tempo fuor di sesto**  
( *Time Out of Joint* )  
Philip K. Dick  
2003  
2006  
2007

# Il Dataset:

Limitazione al  
sottogenere  
Fantascienza/Cyberpunk  
e agli autori collegati

Web scraper realizzati  
in *Python*  
(librerie *BeautifulSoup* e  
*MechanicalSoup*)

**Catalogo Vegetti della letteratura fantastica**  
BETA PUBBLICATA IN ITALIA - A CURA DELLA GILDA DEI CATALOGATORI

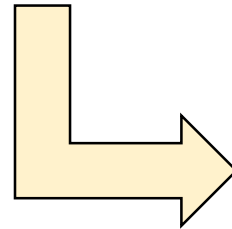
*Scraping di titolo, autore, copertina, codice NILF (chiave univoca)*



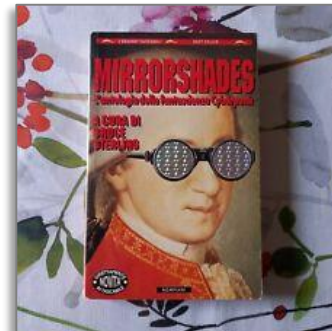
*Scraping di fotografie*



*Scraping aggiuntivo (libri con pochi campioni)*



**NILF\_DB**  
*656 volumi*  
*4839 fotografie*



# Augmentation:

Immagini sottoposte a  
scaling e padding

## Split:

Training set	70%
Validation set	20%
Test set	10%

## Trasformazioni:

- Luminosità e contrasto
- Sfocatura e nitidezza
- Traslazione, rotazione e scala
- Shear (leggero)
- Ricolorazione (leggera)
- Rumore gaussiano (aggiunto)





# Le Reti Neurali:

## Addestrate da zero:

- *LeNet-5*
- Paper “*Judging a Book by its Cover*”
- Paper “*Deep Learning Approaches towards Book Covers Classification*”

Risultati insoddisfacenti con reti addestrate da zero

## Transfer learning:

- *AlexNet*
- *ResNet*
- *Inception*

Ottimi risultati in *PyTorch* con reti pre-addestrate su *ImageNet*

Quantization Aware  
Training  
+ ottimizzazione ARM

Nome del modello	Dimensione del campione	Accuracy (fast)	Accuracy (full)	Note
AlexNet	$227 \times 227 \times 3$	84,4 %	86,4 %	storage richiesto > 200 MB
ResNet-18	$224 \times 224 \times 3$	96,0 %	98,9 %	
ResNet-50	$224 \times 224 \times 3$	96,7 %	-	
Inception-v3	$299 \times 299 \times 3$	89,7 %	-	

Modello quantizzato	Dimensione del campione	Accuracy (fast)	Accuracy (full)	Note
AlexNet	$227 \times 227 \times 3$	85,6 %	86,8 %	PTQ (QAT non supportata)
ResNet-18	$224 \times 224 \times 3$	89,1 %	98,0 %	
ResNet-50	$224 \times 224 \times 3$	91,6 %	-	tempo di inferenza > 10 s
Inception-v3	$299 \times 299 \times 3$	-	-	TorchScript non supportato



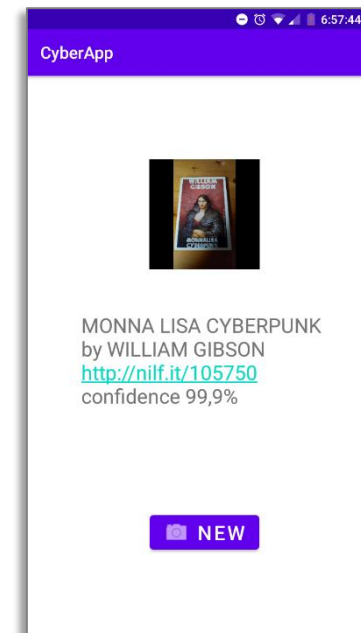
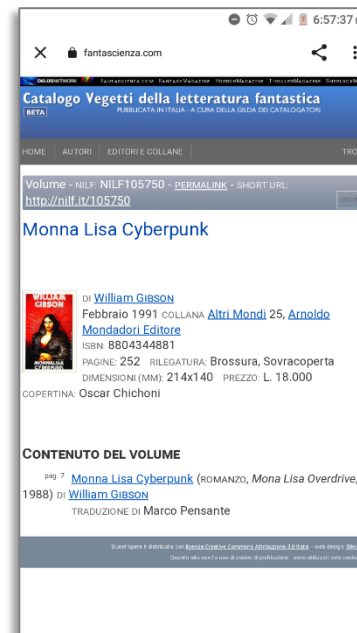
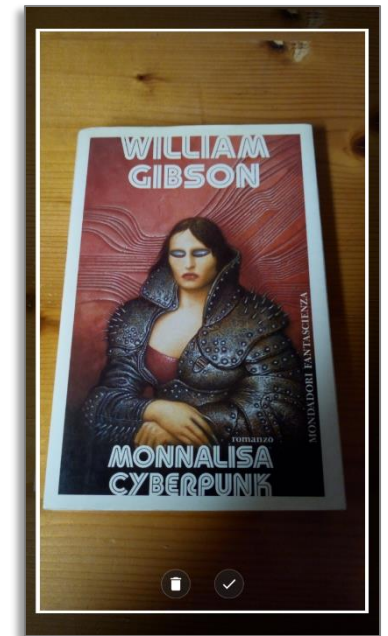
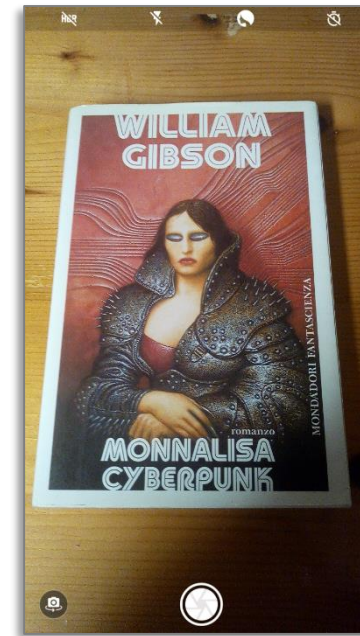
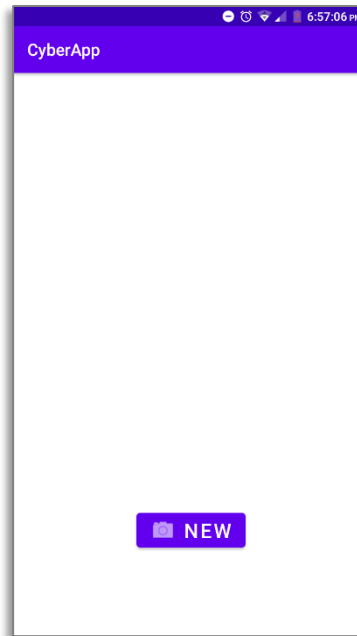
La rete selezionata per il deployment è la ***ResNet-18 (full)***

# L'Applicazione:

*Kotlin + PyTorch Mobile*

Include la rete e le coppie  
titolo-autore da associare  
all'etichetta predetta

Il link alla pagina del  
Catalogo è derivato  
dall'etichetta



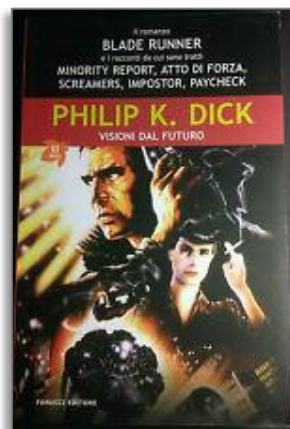
# Conclusioni:

Accuratezza soddisfacente:  
98,0%

Alcuni fallimenti degni di nota

## Sviluppi futuri:

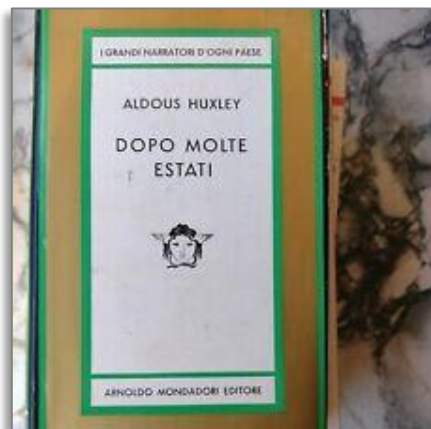
- Estendere il dataset ripetendo la ricerca di immagini
- Semplificare la pulizia manuale introducendo altri classificatori



*Stessa illustrazione molto dettagliata*



*Pochi esempi e di scarsa qualità*



*Collana non illustrata e stesso autore*



*Stessa collana e stesso illustratore*



*...ma anche molti casi di successo!*

**Grazie.**

*Codice e dataset:*

<https://github.com/JackFantaz/CyberApp>