

Proje Başlığı:

İlaç Kutularından Yazı Tanıma ve İlaç
Bilgisi Çıkarımı İçin YOLO ve CNN
Tabanlı Hibrit Bir Yaklaşım

Hazırlayan: Orhancan Yıldırım

Tarih: 10/05/2025

Özet

Bu çalışma, ilaç kutuları üzerindeki metinleri tanımak ve tanınan metne uygun ilaç bilgilerini veri tabanından çekmek için YOLOv8 tabanlı bir kutu tespiti modeli ve CNN tabanlı bir OCR (Optik Karakter Tanıma) modelinin entegre kullanıldığı bir sistem sunmaktadır. Geliştirilen sistem, görseldeki metni önce YOLOv8 ile lokalize eder, ardından lokalize edilen kutu ya da tüm görsel üzerinden CNN tabanlı OCR işlemi yaparak kelime tahmini gerçekleştirir. Elde edilen kelime, RapidFuzz kütüphanesi aracılığıyla bir CSV veri kümesinde aranır ve eşleşen ilaç kaydının bilgileri kullanıcıya sunulur. Sistem React tabanlı bir kullanıcı arayüzü üzerinden kolay kullanılabilir bir formda sunulmaktadır. Yapılan testler sistemin yüksek doğrulukta tahmin yapabildiğini ve ilaç veri kümeleri ile başarılı eşleşmeler sağladığını göstermiştir.

Anahtar Kelimeler: OCR, YOLO, CNN, İlaç Tespiti, RapidFuzz

Giriş

Sağlık teknolojilerinde görsel tabanlı akıllı sistemler büyük önem kazanmıştır. Özellikle ilaç kutuları üzerinde yer alan metinlerin doğru bir şekilde okunarak ilaç hakkında bilgiye ulaşılması; eczacılık, sağlık yönetimi ve hasta güvenliği açısından kritik rol oynamaktadır.

Klasik OCR sistemleri sabit yazı formatlarıyla sınırlıdır ve gerçek dünya koşullarında (farklı fontlar, ışık, ambalaj karmaşası) yetersiz kalmaktadır. Bu proje kapsamında, ilaç kutularındaki yazıları hem tespit eden hem de okuyan akıllı bir sistem geliştirilmiştir. Sistem YOLOv8 kullanarak kutudaki metni lokalize etmekte ve CNN tabanlı bir model ile karakter tanıyarak kelime oluşturmaktadır.

Literatürde tıbbi belgeler veya reçete OCR sistemleri üzerine birçok çalışma yapılmış olsa da, ambalaj üzerindeki yazıların okunması daha az çalışılmıştır. Bu proje, bu boşluğu doldurmak amacıyla geliştirilmiştir.

4. Yöntem (Metodoloji)

Bu projede hibrit bir yaklaşım uygulanmıştır:

4.1 Kutu Tespiti - YOLOv8

- Eğitilmiş YOLOv8 modeli, ilaç kutusu görselinde metin içeren alanları tespit eder.
- Eşik değeri: 0.4 (güven skoru)
- Eğer kutu bulunamazsa tüm görsel OCR işlemine gönderilir.

4.2 OCR Aşaması - CNN

- PyTorch ile eğitilen OCRCNN modeli kullanılmıştır.
- Ön işleme: gri tonlama, histogram eşitleme, normalize, yeniden boyutlandırma
- Giriş görseli: 32x32 piksel, tek kanal

4.3 Kelime Seçimi

- Görselde bulunan karakter konturları yüksekliğe göre gruplanır.
- Aynı satıra ait karakterler gruplanır ve ortalama yükseklikleri alınır.
- En yüksek ortalama yüksekliğe sahip satır, asıl yazı satırı olarak kabul edilir.

4.4 Eşleştirme

- OCR sonucu olan kelime, medicine_info.csv dosyasında RapidFuzz ile aranır.
- Benzerlik skoru ≥ 70 olan ilk eşleşme alınır.

4.5 Araçlar

- FastAPI (backend)
- React (frontend)
- OpenCV, PyTorch, Ultralytics YOLOv8, Pandas, RapidFuzz

5. Bulgular ve Tartışma

YOLOv8 Kutu Tespiti Sonuçları:

- mAP50: %92.6
- mAP50-95: %86.9
- Precision: 0.85, Recall: 0.82

OCR Modeli (OCRCNN) Sonuçları:

- Maksimum test doğruluğu: %92.8
- Kayıp fonksiyonu eğitim sürecinde azalmıştır (0.42 \rightarrow 0.25)

Tartışma

- YOLOv8 ile yüksek doğrulukta kutu tespiti sağlanmıştır.
- Kutu tespiti yapılamayan durumlarda en büyük yazıyı seçme stratejisiyle hatalar azaltılmıştır.
- Sistem Parol, Napa gibi örneklerde yüksek başarı sağlamıştır.

6. Sonuç ve Öneriler

Bu proje, ilaç kutularının görsellerinden metin tanıyarak ilgili ilaç bilgisini çıkarmayı başaran bir sistem ortaya koymuştur. YOLO ve CNN gibi güçlü derin öğrenme modellerinin entegrasyonu sayesinde, sistem gerçek dünya koşullarında da işlevsel şekilde çalışmaktadır.

Sistem farklı veri kümeleri, çok dilli içerikler ve mobil cihazlara uyumlu hale getirilerek yaygınlaştırılabilir. Ayrıca barkod/QR kod desteği, segmentasyon ve dil modeli düzeltmesi gibi eklemelerle doğruluk daha da artırılabilir.

7. Kaynakça

- <https://www.kaggle.com/datasets/toriqulstu/all-medicine-and-drug-price-data20k-bangladesh>
- <https://www.kaggle.com/datasets/harieh/ocr-dataset>
- <https://www.kaggle.com/datasets/preatcher/standard-ocr-dataset>
- <https://maxbachmann.github.io/RapidFuzz/>
- <https://www.kaggle.com/code/nikhilteja21/character-recognition-cnn>
- <https://www.kaggle.com/code/harieh/cnn-model-implementing-ocr>