

Jak nauczyć maszynę identyfikować i zliczać obiekty w piekarni

– wykorzystanie Computer Vision pakietu Azure Cognitive Services



Przypomnijmy sobie słynną scenę z filmu *Rain Man*¹, w której bohaterowie znajdują się w restauracji, a na podłogę przypadkowo wysypują się wykałaczki. Raymond, grany przez Dustina Hoffmana, spogląda przez chwilę na rozsypane wykałaczki i bez chwili namysłu podaje dokładną ich liczbę – 246. Kasjerka chwilę później potwierdza, że opakowanie zawierało 250 patyczków, z czego 4 pozostały w pudełku. Dla przeciętnego człowieka taka zdolność wydaje się wręcz nadludzka – to przykład tzw. sawantyzmu, czyli niezwykle rozwiniętych umiejętności wąskiego zakresu u osób z zaburzeniami neurorozwojowymi.



A teraz wyobraźmy sobie podobną scenę w naszym zakładzie produkcyjnym. Zamiast wykałaczek mamy bułki, chleby, ciasteczka, opakowania. Zamiast geniusza z sawantyzmem – komputer z dostępem do kamery i sztuczną inteligencją.

A co by było, gdybyśmy komputer nauczili w ten sposób identyfikować ilości naszych produktów? Czy mógłby w mgnieniu oka zliczyć wyroby na linii, zweryfikować kompletację zamówienia, sprawdzić, czy żaden koszyk nie został pomyłony? Technologia już dziś daje nam taką możliwość – wystarczy ją odpowiednio wdrożyć.

¹ Film *Rain Man* to amerykański dramat z 1988 roku w reżyserii Barry'ego Levinsona, który zdobył aż cztery Oscary, w tym za najlepszy film, najlepszego reżysera, najlepszy scenariusz oryginalny i najlepszego aktora pierwszoplanowego – dla Dustina Hoffmana.

Kompletacja w piekarni – wolimy nie pamiętać, że jest

W nowoczesnych piekarniach i cukierniach proces kompletacji stanowi jeden z najbardziej wymagających etapów całej działalności firmy. Choć często niedostrzegany z zewnątrz, to właśnie on w dużej mierze determinuje skuteczność dostaw. Kompletacja, czyli tworzenie zestawów produktów dla poszczególnych punktów sprzedaży, odbywa się w warunkach dużej presji czasowej – najczęściej wczesnym rankiem, tuż po zakończeniu nocnej produkcji.

W tym niewralgicznym momencie każda jednostka pieczywa, każde ciasto i każda partia ciastek muszą zostać przyporządkowane odpowiednim sklepom, posegregowane zgodnie z tabelami przydziału i bezbłędnie

załadowane do pojazdów, których trasy zostały wcześniej zaplanowane. Całość stanowi misternie zaprojektowany proces logistyczny, który – jak się okazuje – nie rzadko przewyższa swoją złożonością sam etap wypieku.

Dodatkowym utrudnieniem są warunki, w jakich realizowana jest kompletacja: ograniczona widoczność, niskie temperatury oraz obniżona sprawność poznawcza pracowników ze względu na porę dnia. Wrażliwe produkty, takie jak wyroby cukiernicze, wymagają ponadto zastosowania specjalistycznych opakowań ochronnych, co jeszcze bardziej podnosi wymagania organizacyjne.

Na domiar złego warto wspomnieć o tym, że kompletacji dokonują osoby, które nie należą do osób najlepiej wykształconych i samodzielnych. Nie ma zbyt wielu chętnych żeby wykonywać tę ciężką pracę o tak wczesnej porze i za niską pensję.

W artykule będę starał się przedstawić w jaki sposób za pomocą algorytmów sztucznej inteligencji można zmniejszyć uciążliwość tego procesu.

Niech to zrobią za nas roboty

W obliczu tak dynamicznego i szybkiego rozwoju technologicznego intuicja odpowiada abyśmy chwilę poczekali, a może za jakieś 3 lata wszystkim zajmą się roboty. Otóż żadne roboty nie pomogą, owszem istnieją już takie humanoidalne maszyny, lecz są bardzo drogie i nie opłaca się ich kupować, ponieważ to nie ma żadnego sensu ekonomicznego. Po drugie nawet jakby się pojawiły takie tanie plastikowe roboty, to trzeba je wszystkiego nauczyć. Czyli prawdopodobnie trzeba by zrobić dokładnie to co dzisiaj będę starał się omówić w tym artykule. Po trzecie nawet jakby był nauczony i nawet jakby ich było dużo i były tanie, to póki co przy obecnym stanie rozwoju technologicznego nie zastąpią w procesie kompletacji człowieka, bo to jest trochę za skomplikowane nawet dla LLMa².

Jak zjeść wieloryba?

Ogarnięcie procesu logistycznej dystrybucji pieczywa i produktów cukierniczych jest zadaniem bardzo trudnym dla człowieka, jest to również zadanie trudne dla sztucznej inteligencji. Zadanie to jest tak wielkie jak wieloryb. Czy jest możliwe szybkie ulepszenie procesu kompletacji tak, żebyśmy mogli powierzyć je sztucznej inteligencji? Owszem jest, tylko jak każde wielkie dzieło musi być wykonane w drobnych krokach. To cała mozaika małych projektów, które do siebie pasują jak puzzle. Jeżeli będziemy efektywni w realizacji tych małych projektów, to przyjdzie taki dzień, że połączymy wszystkie te drobne

ulepszenia i stworzymy całkowicie autonomiczny system dystrybucji i to praktycznie bez kosztów. Czy obecnym stanie rozwoju technologicznego jest to możliwe? Proces ten przypomina zjadanie wieloryba. W jaki sposób zjeść wieloryba? Trzeba go podzielić na małe części, a następnie codziennie sukcesywnie je zjadać.

Jakiś czas temu napisałem artykuł do tego magazynu³ opisujący w jaki sposób sztuczna inteligencja może zarządzać procesem logistycznym w planowaniu tras samochodów i szybkim wyszukiwaniu rozwiązań alternatywnych.

Opisałem tutaj mały projekt, bardzo łatwy do zrealizowania. W tym przypadku zarządzanie flotą przez sztuczną inteligencję jest usługą, która dostarczana jest przez wyspecjalizowane firmy, tak jak opisałem to we wspomnianym artykule. A ponieważ jest to usługa, musi być tak skonstruowana, żebym była bardzo łatwa dla jej użytkowników. Właściwie nic nie trzeba tutaj robić tylko po prostu trzeba to wdrożyć.

Jak nauczyć kota liczyć?

Kolejnym puzzlem we wdrożeniu sztucznej inteligencji do piekarni mogłaby być analiza wizualna produktów. Żeby uruchomić robota, który mógłby cokolwiek robić, robot musi widzieć i rozpoznawać przedmioty. Żeby przyspieszyć proces kompletacji albo kontroli skompletowanych produktów trzeba by nauczyć jakąś maszynę rozróżniać, zliczać, wykrywać braki albo uszkodzenia produktów.

Samodzielne zrobienie takiej mądrej, liczącej kamery komputerowej jest możliwe tylko jest dość żmudne i czasochłonne. Przy użyciu specjalnych bibliotek w języku Python (Keras czy Pytorch) można nauczyć komputer rozróżniać poszczególne kształty, zliczać kosze itp. Na szczęście uczenie komputera rozróżniać przedmioty jest już usługą na sprzedaż, za którą nie trzeba dużo zapłacić. Dzięki temu cały proces nauki maszyny kontroli kompletacji może okazać się stosunkowo prosty i przyjemny.

Wprowadzenie sztucznej inteligencji do piekarni powinniśmy rozpatrywać w kategoriach małych projektów, a każdy projekt powinien być wykonany w kilka dni. Przede wszystkim musi być zamknięty, musi działać i przynosić konkretne udogodnienia. Taki proces raczej polubić, aby był efektywny.



² LLM to skrót od Large Language Model, czyli duży model językowy. Jest to zaawansowany system sztucznej inteligencji oparty na architekturze sieci neuronowych (najczęściej transformerów), który został wytrenowany na ogromnych zbiorach danych tekstowych.

³ „Zagadnienie transportowe w piekarni”, PRZEGŁĄD PIEKARSKI I CUKIERNICZY | Rocznik 2024 – zeszyt 12

Jak miałyby to wyglądać?

Wyobraźmy sobie, że na rampie załadunkowej stawiamy pojedynczy zestaw dla konkretnego sklepu. Naciskamy przycisk, cztery kamery robią zdjęcia z czterech stron. Następnie ze ściany wychodzi wydruk albo na monitorze pojawia się tabela, w której został zliczony cały asortyment, porównany w tabeli ze stanem oczekiwanym. W ten sposób zostało dokonane sprawdzenie partii do wysyłki. Jeżeli to jest za bardzo wydumany sposób kontroli, to możemy identyczne podejście zastosować instalując kamerę zliczającą ilość pieczywa wychodzącą z poszczególnych pieców. Możemy zliczać produkty, które przysły ze zwrotów i porównywać te wyniki z dokumentami. Możemy też zliczać butki na taśmach produkcyjnych albo dokonywać automatycznych inwentaryzacji w magazynach surowców sypkich. Nie ma znaczenia w jaki sposób wykorzystamy zdolność maszyn do identyfikacji, zliczania i automatycznego szacowania. Niewątpliwie taka umiejętność byłaby przydatna i stanowiłaby mocne wsparcie dla osób prowadzących piekarnie. Samo zliczanie koszy przed kursem i po kursie stanowiłoby duże ułatwienie. Właściciel miałby tabelę, w której z łatwością można by wychwycić gdzie i kiedy liczba koszy w obiegu przestała się zgadzać. Zresztą taką analizę tabelki koszy z całego miesiąca można dać do policzenia bezpłatnej wersji chatGPT4.5. Zrobi to i wskaże kto zgubił kosze.

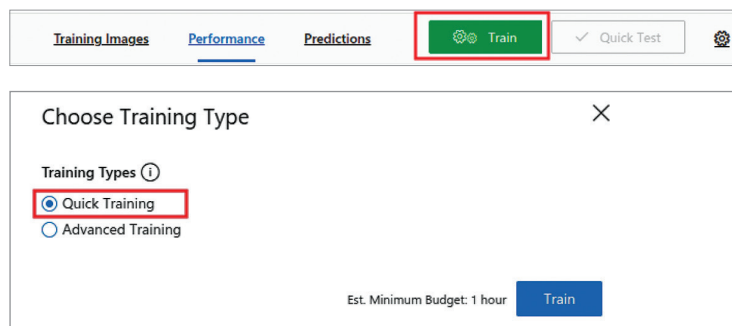
Jak wykrywać ilość chlebów w koszach za pomocą Custom Vision Studio

- Wejść na stronę: <https://customvision.ai>
 - zaloguj się kontem Microsoft, zostaniesz poproszony o połączenie konta z subskrypcją Azure.
- Utwórz nowy projekt
 - kliknij „New Project”
 - uzupełnij:
 - Name: np. kot liczący chleb
 - Description: wykrywanie bochenków chleba w koszach
 - Resource: wybierz istniejący zasób Azure lub utwórz nowy
 - Project Type: Object Detection
 - Domain: najlepiej General [A1]: uniwersalna domena dla większości zastosowań. Wybierz tę opcję, jeśli nie masz specyficznych wymagań dotyczących rodzaju obrazów.
- Dodaj zdjęcia z bochenkami chleba
 - kliknij „Add images”. Wgraj zdjęcia przedstawiające kosze z własnymi bochenkami czy tym co chcesz aby maszyna zliczała
 - każde zdjęcie powinno zawierać od 1 do wielu chlebów w różnych pozycjach i warunkach oświetleniowych.

4. Otaguj zdjęcia (kluczowy krok!)

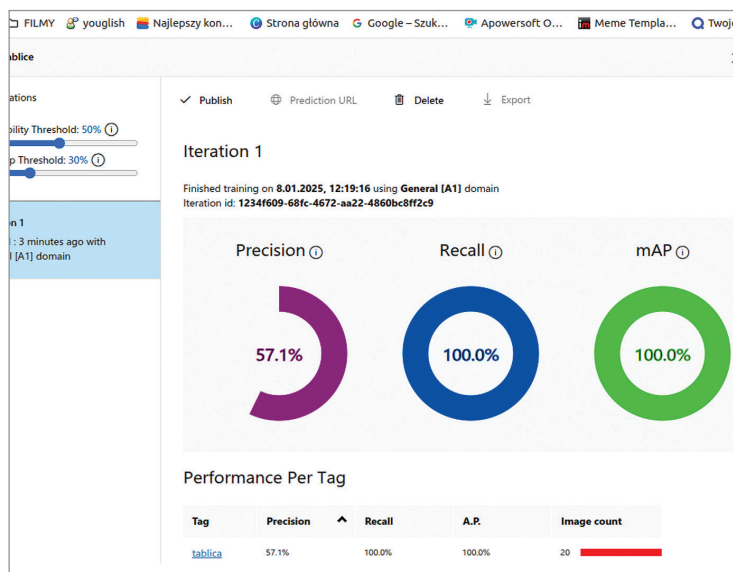
- po dodaniu zdjęć, kliknij zdjęcie. Zaznacz każdy bochenek chleba ramką i nadaj mu tag, np. bread. Im więcej różnych przypadków, tym lepiej działa model. Musi być minimum 30 zdjęć z różnymi ułożeniami chleba.

- Kliknij „Train”
 - wybierz Quick training
 - trening potrwa kilka minut
 - po zakończeniu zobaczysz metryki: Precision, Recall, Average Precision



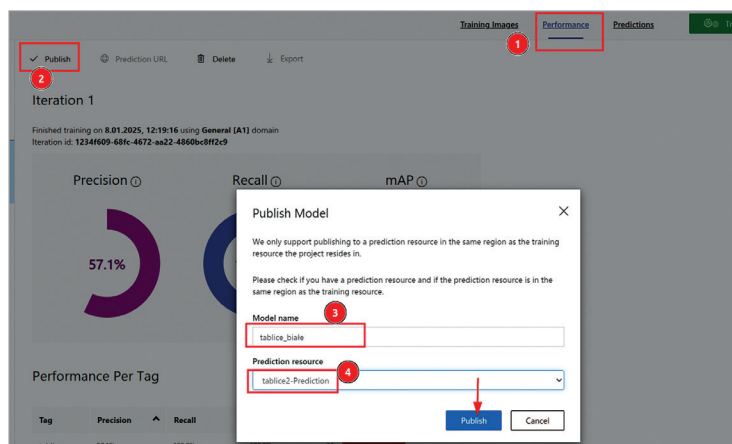
6. Przetestuj swój model

- kliknij „Quick Test”
- prześlij nowe zdjęcie z koszem i sprawdź czy model poprawnie wykrywa i zlicza chleby.

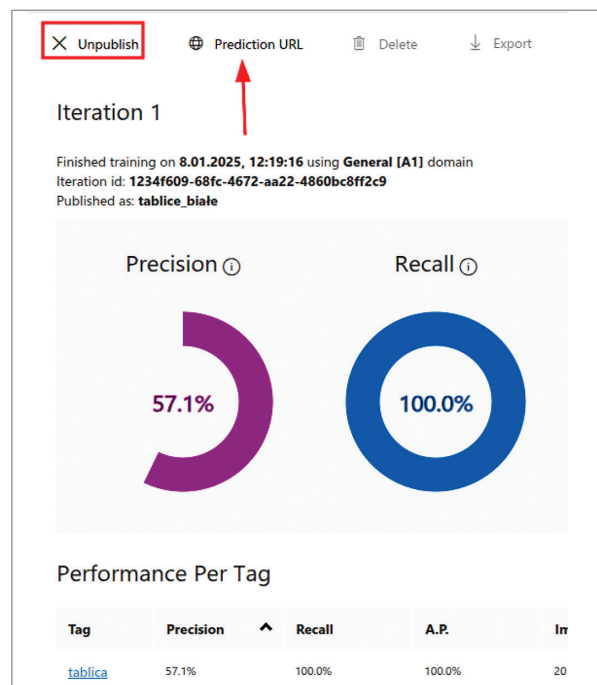


Powyższe wyniki walidacji modelu w parametrze „Precision” mówią, że model zliczył 57% dobrych bochenków, czyli 43% było policzonych jako bochenki, a to były np. bułki lub ciastka. Niestety tutaj troszkę trzeba się pobawić, aby mieć dobre wyniki. Po prostu trzeba dać więcej zdjęć. System będzie działał, gdy zdjęcia do treningu będą podobne do zdjęć w trakcie normalnej pracy modelu. Upewnij się, że obrazy są zróżnicowane pod względem: Kątów kamery, oświetlenia, tła.

Gdy będziemy już zadowoleni z wyników, musimy opublikować model.



Po opublikowaniu modelu dostajemy klucze API dzięki którym możemy połączyć model z lokalnym komputerem, smartfonem lub tabletem.



How to use the Prediction API

If you have an image URL:

```
https://tablica2-prediction.cognitiveservices.azure.com/customvision/v3.0/Predictor
Set Prediction-Key Header to:
E4owLlb6z7ms8iHy4Wgwpv0i5CeZOVnGRvNt6f0dtq3H2EIjT1lqJQQJ99BAAC15YpXJ3W
Set Content-Type Header to: application/json
Set Body to: {"Uri": "https://example.com/image.png"}
```

If you have an image file:

```
https://tablica2-prediction.cognitiveservices.azure.com/customvision/v3.0/Predictor
Set Prediction-Key Header to:
E4owLlb6z7ms8iHy4Wgwpv0i5CeZOVnGRvNt6f0dtq3H2EIjT1lqJQQJ99BAAC15YpXJ3W
Set Content-Type Header to: application/octet-stream
Set Body to: <image file>
```

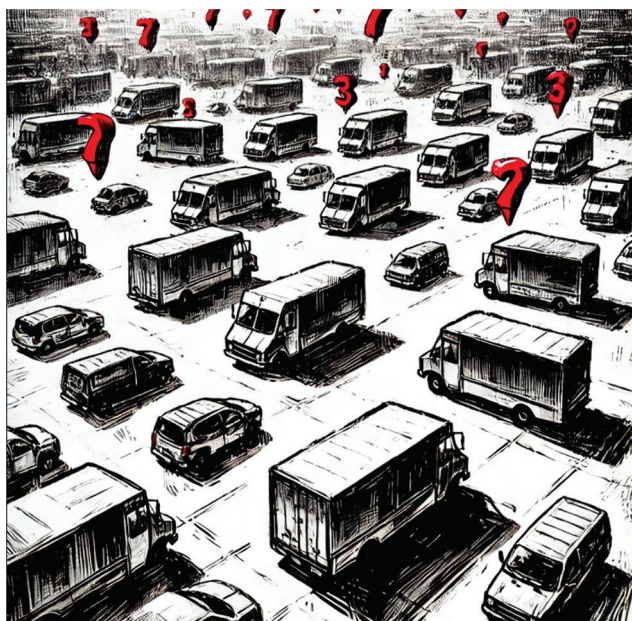
Got it!

Na koniec, co istotne zawsze w biznesie, koszt – jest raczej trywialny w stosunku do efektywności tej usługi.

Cena	
Liczba bezpłatnych transakcji na miesiąc:	5 000
Liczba transakcji na minutę:	20
Bezpłatnie	
Tag	0-1 mln transakcji - \$1 za 1000 transakcji
Rozpoznawanie twarzy	GetThumbnail
Kolor	1-10 mln transakcji - \$0,65 za 1000 transakcji
Typ obrazu	GetAreaOfInterest
Wykrywanie osób (wersja zapoznawcza)	10-100 mln transakcji - \$0,60 za 1000 transakcji
Inteligentne pzyznania	Ponad 100 mln transakcji - \$0,40 za 1000 transakcji
OCR	

Podsumowanie

Przedstawiony opis jest dość enigmatyczny. W publikacji wskazuję w jaki sposób należy kamerę nauczyć liczyć. Oczywiście za pierwszym razem trzeba będzie jeszcze się pomęczyć, aby wszystko wyszło. Za każdym kolejnym razem będzie coraz łatwiej. Załóżmy, że poświęcimy na tę pracę jeden/dwa dni. W rezultacie otrzymujemy efektywne narzędzie, które będzie coś liczyło, np. samochody na parkingu. Załóżmy, że co godzinę będziemy zliczać samochody na parkingu. Maszyna będzie to robiła automatycznie tworząc tabelę czasową.



Będzie tam data godzina i minuta oraz liczba pojazdów. Można też nauczyć komputer, aby rozróżniał te pojazdy. Maszyna może również zliczać klientów wchodzących do sklepu, może też liczyć osoby znajdujące się w kawiarni. Wyobraźmy sobie, że mamy 15 gości, a potem gdy otworzymy rejestr sprzedaży w kasie fiskalnej okaże się, że sprzedano dwie kawy i jednego pączka. A może ktoś sprzedawał coś poza kasą? Automatyczne systemy zliczające mają ogromny potencjał optymalizacji procesów biznesowych. Jak widać z powyższego instrukcji nauczanie maszyny liczyć nie jest procesem przekraczającym możliwości człowieka.

Wojciech Moszczyński

Wojciech Moszczyński – absolwent Katedry Ekonometrii i Statystyki Uniwersytetu Mikołaja Kopernika w Toruniu, specjalista z zakresu ekonometrii, finansów, *data science* oraz rachunkowości zarządczej. Specjalizuje się w optymalizacji procesów produkcyjnych i logistycznych. Prowadzi badania w obszarze rozwoju i zastosowania sztucznej inteligencji. Od lat zaangażowany w popularyzację *machine learning* oraz *data science* w środowiskach biznesowych.