

Dealing with Limited Sample in Office Documents Classification

Paweł Baranowski, Adrian Stępniaś

Instytut Ekspertyz Ekonomicznych i Finansowych oraz Uniwersytet Łódzki

19 Maja 2021, seminarium Katedry Informatyki Stosowanej, AGH

Wprowadzenie - o IEEF

Instytut Ekspertyz Ekonomicznych i Finansowych w Łodzi jest jednostką naukową powołaną w celu sporządzania ekspertyz i opinii na zlecenie sądów i prokuratur (jako dowody z opinii biegłego).

Instytut wydaje opinie głównie w zakresie ekonomii, finansów w tym rynków kapitałowych, bankowości i ubezpieczeń oraz rachunkowości i podatków, szczególnie w sprawach gospodarczych "wielkiej wagi".

Wprowadzenie - motywacja

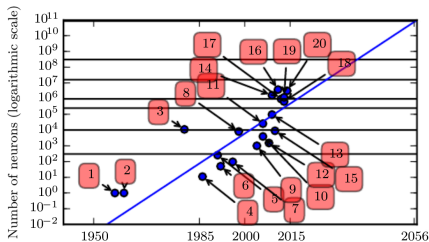
W dużych sprawach gospodarczych akta przekazywane biegłym liczą nawet 100 tys. stron (500 tomów). Udostępniane są (w najlepszym przypadku) w postaci nieustrukturyzowanych skanów.

Tabele są szczególnie ważne dla biegłych z zakresu bankowości i finansów / rachunkowości, gdyż zawierają:

- sprawozdania finansowe
- dane statystyczne
- rejestry księgowe
- wyciągi bankowe
- i wiele innych...

Motywacja c.d.

- Sieci neur. *computer vision* stają się coraz bardziej rozbudowane
- Wraz ze wzrostem złożoności modelu, rośnie rozmiar próby wymaganej do uogólnienia wyników
- "Koszty" ręcznej anotacji zbioru danych w uczeniu nadzorowanym



Goodfellow et al., 2016

Cel badania

Cel: opracowanie sieci neuronowej do automatycznego wykrywania tabel i weryfikacja jej skuteczności w przypadku umiarkowanie licznej próby treningowej.

Dane są wrażliwe - nie możemy użyć istniejących *model as a service*.

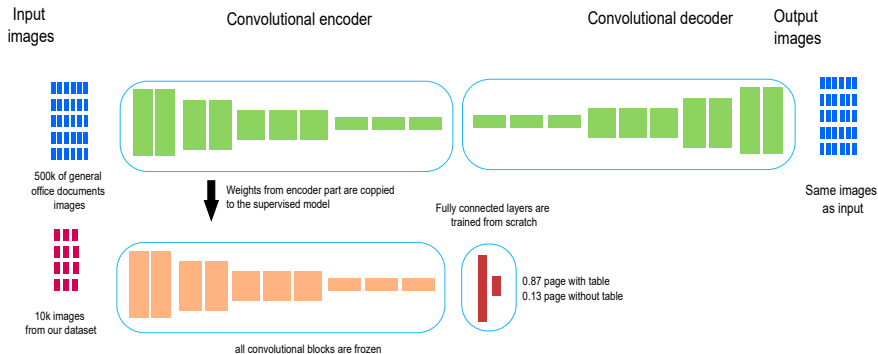
Pytanie: czy przeniesienie parametrów pomiędzy modelami (czysty *transfer learning* lub uczenie częściowo nadzorowane) poprawia zdolności klasyfikacji – wykrywania tabel?

Używamy konwolucyjnej sieci neuronowej przystosowanej do klasyfikacji binarnej (na wyjściu pojedynczy neuron, sigmoidalna funkcja aktywacji).

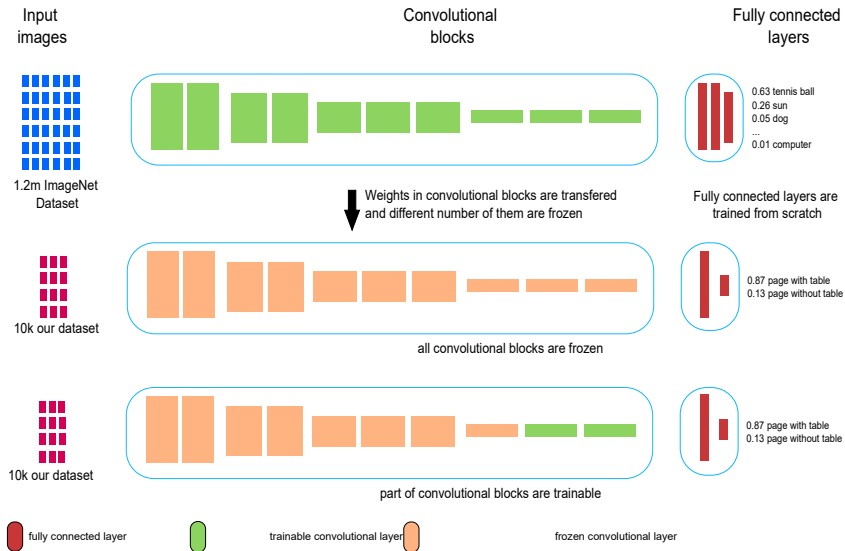
Porównujemy trzy podejścia do uczenia maszynowego:

- Uczenie nadzorowane (supervised learning)
- Uczenie częściowo nadzorowane (semi supervised learning)
- Transfer learning
 - ile warstw konw. przenosimy? (wszystkie / 75% / 50%)

Semi supervised learning



Transfer learning



Zbiór danych

Zbiór danych zawiera ponad 200 tys. stron akt (skany w odcieniach szarości), z czego 11 tys. zostało ręcznie oznaczone flagą binarną "Czy strona zawiera co najmniej jedną tabelę" (32% zawiera tabelę - przykłady).

- losowy podział na próbę treningową i testową (20% == 2190 obs.)
- wzbogacenie-modyfikacja zbioru danych (*data augmentation*): obrót lub pochylenie do 15 st., odbicia lustrzane, powiększenie do 10%

Strona 1 z 1

Wzrost: 170 cm

Waga: 70 kg

Temat: ...

Wzrost: 170 cm

Waga: 70 kg

Temat: ...

Wzrost	Waga	Temat
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...

Strona 2 z 2

Wzrost	Waga	Temat
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...

Strona 3 z 3

Wzrost	Waga	Temat
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...

Strona 4 z 4

Wzrost	Waga	Temat
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...
170	70	...

Architektura modeli

Drugim wymiarem analiz jest architektura sieci neuronowej.

Stosujemy cztery popularne rozwiązania w modelach analizy obrazu.

- VGG16 [link](#)
- Resnet50 [link](#)
- Inception [link](#)
- Xception [link](#)

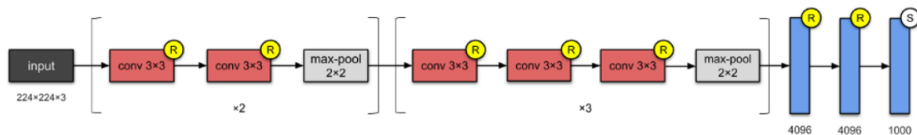
Architektura	Parametrów	Warstw konw.	Gałęzi	Skip-connections
VGG-16	15 mln.	13	1	Nie
ResNet-50 V2	26 mln.	182	do 2	Tak
Inception V3	24 mln.	94	do 4	Nie
Xception	23 mln.	111	do 2	Tak

Supervised	VGG16	Resnet50	Inception	Xception
Accuracy	0.935	0.847	0.867	0.849
Recall	0.940	0.578	0.616	0.573
Transfer (Medium)	VGG16	Resnet50	Inception	Xception
Accuracy	0.966	0.965	0.963	0.973
Recall	0.962	0.922	0.938	0.934
Transfer (Large)	VGG16	Resnet50	Inception	Xception
Accuracy	0.948	0.952	0.948	0.952
Recall	0.936	0.901	0.908	0.894
Transfer (Full)	VGG16	Resnet50	Inception	Xception
Accuracy	0.868	0.931	0.927	0.913
Recall	0.640	0.846	0.855	0.797
Semi-supervised	VGG16	Resnet50	Inception	Xception
Accuracy	0.939	0.801	0.913	0.926
Recall	0.856	0.545	0.774	0.959

- Transfer learning istotnie poprawia zdolność klasyfikacji.
- Szczególnie dobre wyniki dla częściowego transfer learning (50% początkowych warstw konwolucyjnych);
dokładność i czułość $> 95\%$.
- Opracowane modele zostaną wdrożone w codziennej pracy IEEF.
- Zbiór walidacyjny - wyniki zbliżone, z wyjątkiem VGG.

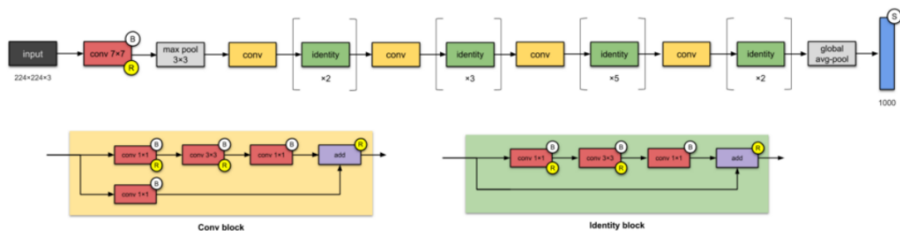
Dziękujemy za uwagę - komentarze lub
pytania?

Załącznik: VGG16 - architektura



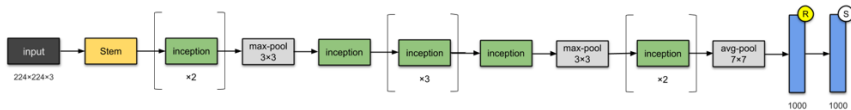
[Powrót](#)

Załącznik: ResNet50 - wizualizacja

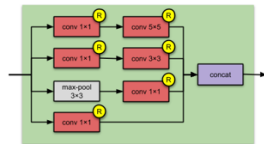


Powrót

Załącznik: Inception - wizualizacja



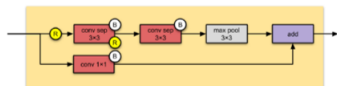
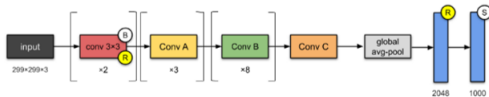
Stem



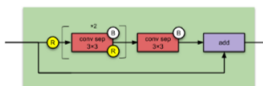
Inception module

Powrót

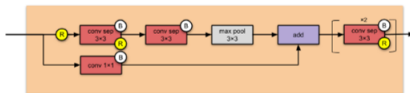
Załącznik: Xception - wizualizacja



Conv A



Conv B



Conv C

[Powrót](#)