

Różnicowanie typów nowotworów nerek na podstawie analizy obrazów tomograficznych

Konkurs Data Science Masters 2020

mgr inż. Aleksandra Osowska-Kurczab
promotor: prof. dr hab. inż. Tomasz Markiewicz

ZETIS, IETISIP
Wydział Elektryczny
Politechnika Warszawska

04.06.2020

mgr inż. Aleksandra Osowska-Kurczab

doktorantka w Szkole Doktorskiej nr 3
Wydział Elektryczny | Politechnika Warszawska

aleksandra.osowska.dokt@pw.edu.pl

Zainteresowania naukowe:

- Computer Vision
- Natural Language Processing
- Zastosowanie technik uczenia maszynowego i głębokiego uczenia do zagadnień medycznych

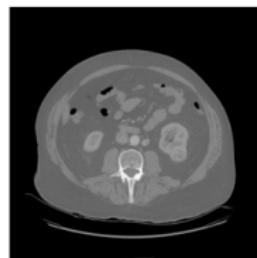


Temat pracy

"Różnicowanie typów nowotworów nerek na podstawie analizy obrazów tomograficznych"

Cel pracy:

- opracowanie algorytmu opisu zmian nowotworowych nerek w obrazowaniu tomografii komputerowej
- przygotowanie zbioru danych, eksploracyjna analiza, normalizacja
- poszukiwanie wzorców typów nowotworów i ich klasyfikacja



Clear Cell Renal
Cell Carcinoma

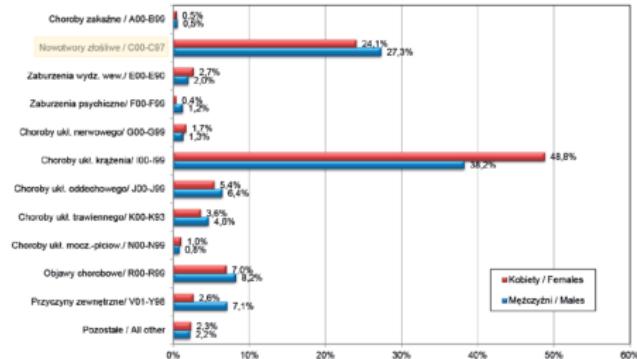
Rysunek 1: Główna idea systemu klasyfikującego

Motywacja

Skąd taki temat pracy?

- system opieki zdrowotnej staje się coraz mniej wydolny (coraz dłuższe kolejki do specjalistów)
- udział chorób nowotworowych w śmiertelności Polaków rośnie
- możliwość leczenia jest silnie powiązana z etapem rozwoju zmiany nowotworowej

Wczesna diagnostyka jest kluczem w obniżeniu statystyk umieralności na choroby nowotworowe.

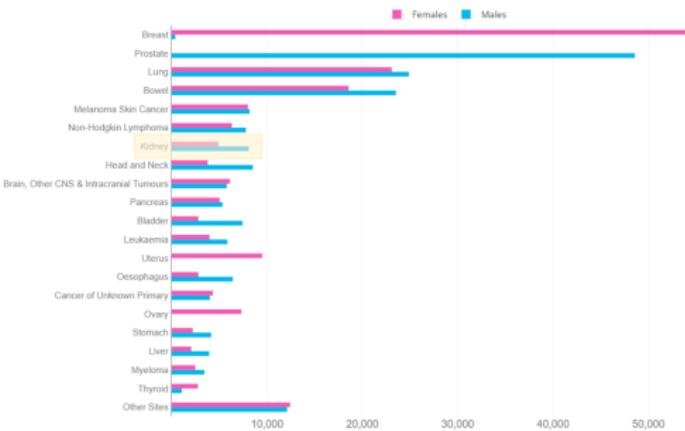


Rysunek 2: Przyczyny zgonów w Polsce wg GUS 2016 [1]

Nowotwór nerki

Charakterystyka nowotworów nerki [3]:

- jeden z rzadszych typów nowotworów
- początkowe stadia choroby przebiegają bezobjawowo
- ponad 50% przypadków wykrywanych jest przypadkowo
- głównie dotyczy osoby w podeszłym wieku



Rysunek 3: Liczba rozpoznanych przypadków nowotworów w UK 2017 [2]



Metody przedstawiane w literaturze

Temat rozpoznawania typów nowotworów nerek nie jest tematem bardzo popularnym. Prac naukowych bezpośrednio poświęconych automatycznej analizie badań CT pacjentów z nowotworem nerki jest bardzo niewiele.

Jakie metody można napotkać w literaturze?

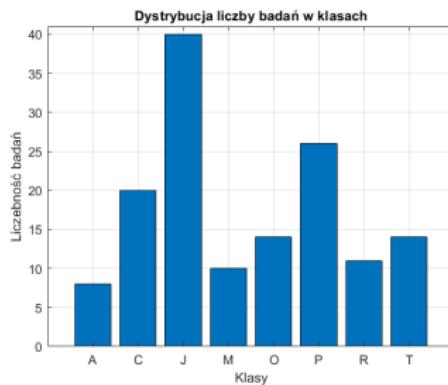
- deskryptory GLCM + SVM/KNN [4]
- filtracja Gabora [5]
- kaskady klasyfikatorów binarnych [6]
- zespoły klasyfikatorów
- CTTA [7, 8]
- 3D-CNN AlexNet/GoogleNet, sieci hybrydowe 3D-CNN + SVM [9, 10]
- DBN, autoenkodery [11]
- segmentacja: U-Net [12], V-Net, metody morfologiczne

Zbiór danych

Na potrzebę badań w grancie OPUS 12 nr 2016/23/BSt6/00621 zgromadzony został zbiór 143 badań tomografii komputerowej pacjentów z 8 typami nowotworów nerek. Dzięki współpracy z lekarzami z Wojskowego Instytutu Medycznego w Warszawie możliwa była anotacja badań oraz analiza wyników badań histopatologicznych.

Tabela 1: Liczebność badań oraz ramek

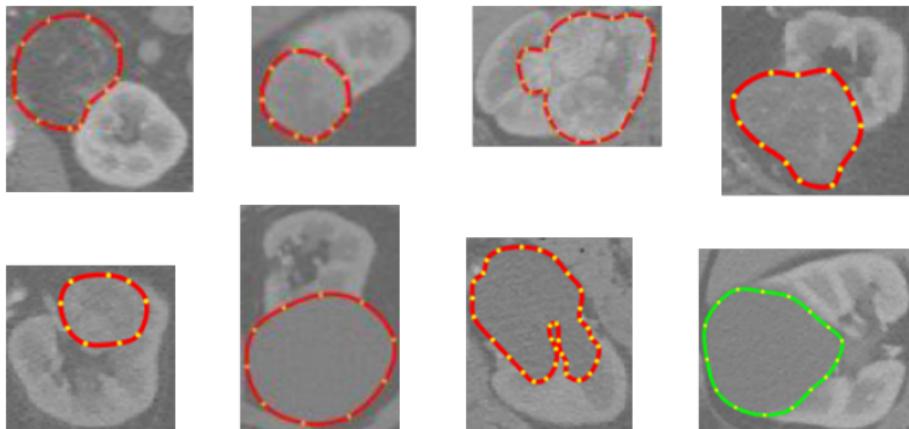
| Pełna nazwa | Ozn. | Liczba badań | Liczba klatek badania |
|--|------|--------------|-----------------------|
| Angiomolipoma | A | 8 | 97 |
| Chromophobe Renal Cell Carcinoma | C | 20 | 253 |
| Clear Cell Renal Cell Carcinoma | J | 40 | 692 |
| Multilocular Cystic Renal Cell Carcinoma | M | 10 | 164 |
| Oncocytoma | O | 14 | 108 |
| Papillary Renal Cell Carcinoma | P | 26 | 236 |
| Urothelial Carcinoma | R | 11 | 292 |
| Renal Cystis | T | 14 | 460 |



Rysunek 4: Rozkład liczby badań (pacjentów) w podziale na klasy



Przykłady klas

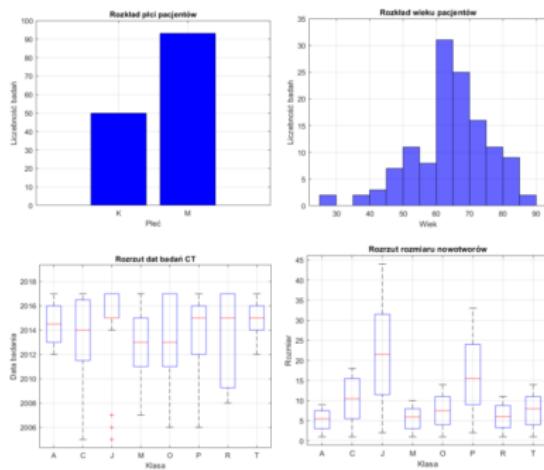


Rysunek 5: Przykładowe fragmenty ramek badania CT każdej z klas, zawierające obszar nerki z obrysowanym nowotworem (od lewej: A, C, J, M, poniżej: O, P, R, T).

Eksploracyjna analiza zbioru

Cechy zgromadzonego zbioru:

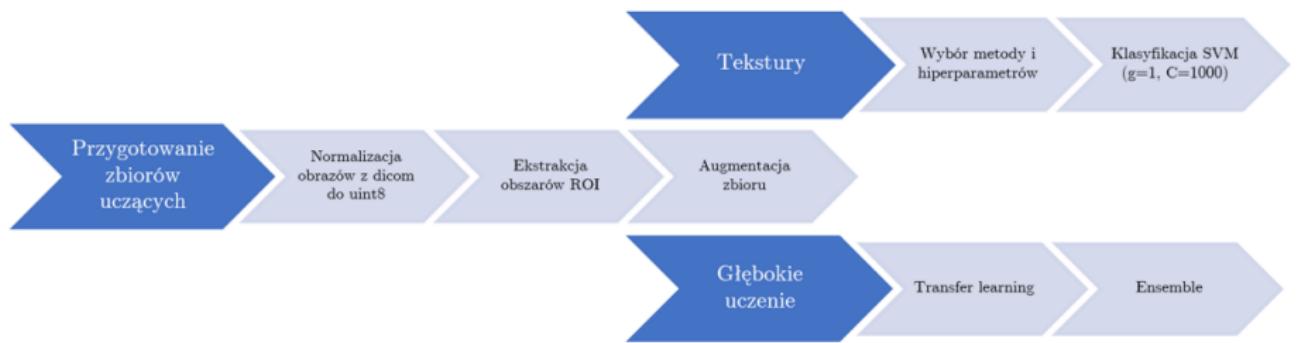
- niezbalansowanie klas
- mały rozmiar nowotworu
- duża różnorodność wewnętrzna klas
- obraz 1 kanałowy
- niski kontrast
- różne tomografy
- zróżnicowana rozdzielczość



Rysunek 6: Eksploracyjna analiza zbioru



Architektura systemu



Rysunek 7: Architektura i operacje wykorzystywane w zaproponowanym systemie

Przygotowanie danych



① **Normalizacja** zbioru wartości obrazów do uint8

② **Ekstrakcja ROI** - 5 możliwości (tabela 2)

③ **Augmentacja** - algorytm wykorzystujący rotację i przycinanie na największych obrazach. Cel: wyrównanie liczebności klas przy jak najmniejszym zniekształceniu pierwotnego obrazu.

Tabela 2: Zaproponowane typy obszarów ROI

| Pierwotna ramka | Typ | ROI |
|-----------------|-----|-----|
| | D1 | |
| | D2 | |
| | D3 | |
| | D4 | |
| | D5 | |



Metody teksturalne - charakteryzują liczbowo regiony obrazu zawierające obszary o właściwościach takich jak szorstkość, nierówność i gładkość. Opis tekstury jest funkcją intensywności pikseli obrazu.

- GLCM z wektorem Haralick'a [13]
- algorytmy fraktalne (SFTA) [14, 12]
- cechy Unser'a [15, 16]
- filtracja Gabora
- LBP [17], HOG [18], cechy statystyczne 2D

Po etapie generacji cech nastąpiła klasyfikacja z użyciem **SVM** [19] o jądrze gaussowskim z parametrami $g=1$, $C=1000$. Zarówno w analizie teksturalnej jak i sieciach głębokich do badania skuteczności systemu wykorzystywane były klasyczne metryki klasyfikacji (Acc, Prec, Rec, **F1** [20]) oraz *confusion matrix*.

Testowano również wpływ redukcji **PCA** na końcowe wyniki, wizualizowano przestrzeń cech z użyciem **TSNE** oraz generowano **macierz Fishera** określającą istotność cech.



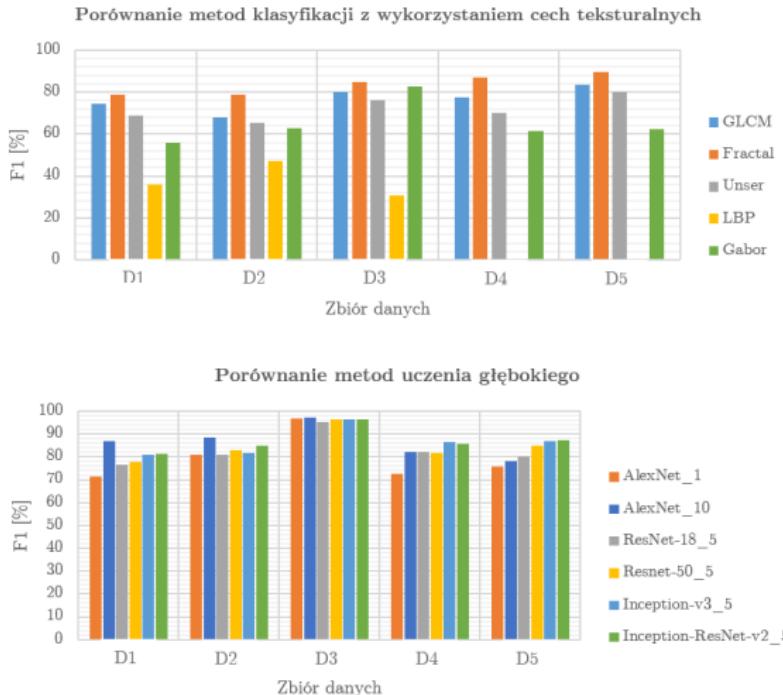
Wykorzystywana była technika **transfer learning'u**[21, 22, 23] popularnych modeli pretrenowanych do klasyfikacji ImageNet. Została ona wybrana ze względu na bardzo ograniczoną liczebność zgromadzonego zbioru. Modele trenowane od losowych wartości nie były w stanie przekroczyć 60% F1.

Testowane architektury:

- AlexNet [24]
- ResNet-18, -50 [25]
- Inception-v3 [26, 27]
- InceptionResNet-v2 [28]

Badana była zależność skuteczności klasyfikacji w zależności od przyjętej warstwy odcięcia (w technice transfer learning) oraz wynikowe prawdopodobieństwo (pewność wskazania). Dotrenowane sieci stwarzyszano w zespół (**ensemble**), którego wskazaniem było głosowanie większościowe członków. Sprawdzano również **modele hybrydowe** (aktywacje sieci pretrenowanych + SVM).

Wyniki



Rysunek 8: Zbiorcze wyniki F1 badań w podziale na typy generacji zbioru



Dyskusja

Tabela 3: Najlepsze rezultaty eksperymentów analizy teksturalnej [w %, 10-fold cross validation]

| Nr | Metoda | Typ ROI | Acc | Prec | Rec | F1 |
|----|---------|---------|------|------|------|------|
| 1 | Fractal | D5 | 89.0 | 89.7 | 89.0 | 89.1 |
| 2 | Fractal | D5 | 88.5 | 88.8 | 88.5 | 88.5 |
| 3 | GLCM | D5 | 80.9 | 81.5 | 81.1 | 81.0 |

Metody teksturalne:

- najlepsze wyniki uzyskano dla metody fraktalnej (89.1% F1) dla zbioru D5
- tylko część testowanych metod nadaje się do wykorzystania w zadaniu różnicowania typów nowotworów nerek
- błędy popełniane przez system oparty o analizę teksturalną przypominają błędy popełniane przez człowieka
- ważnym etapem może okazać się *feature selection* oraz dopasowywanie hiperparametrów metod (np. definicję przyjętego otoczenia w GLCM)
- być może połączenie wielu typów cech pozwoli polepszyć końcową skuteczność systemu



Tabela 4: Najlepsze rezultaty eksperymentów z uczeniem głębokim
[w %, 10-fold cross validation]

| Nr | Metoda | Typ ROI | Acc | Prec | Rec | F1 |
|----|-----------------------|---------|------|------|------|------|
| 1 | AlexNet 10 | D2 | 87.1 | 87.5 | 87.1 | 87.1 |
| 2 | Inception-ResNet-v2 5 | D5 | 86.4 | 86.7 | 86.4 | 86.5 |
| 3 | Inception-v3 5 | D4 | 85.7 | 86.0 | 85.3 | 85.7 |

Transfer learning:

- najlepsze wyniki uzyskano dla zespołu 10 sieci AlexNet (87.1% F1) dla typu generacji D2
- zastosowanie techniki ensemble potrafi poprawić wyniki aż o 8pp względem pojedynczego członka zespołu
- modele pretrenowane (AlexNet, ResNet, Inception) uzyskują porównywalne wyniki skuteczności
- technika transfer learning'u jest wygodna z punktu widzenia projektanta systemu - zastosowanie sieci CNN powoduje, że znika etap żmudnego generowania cech (*feature engineering*)
- dużym problemem w osiąganiu wyższych skuteczności może być resizing obrazów o rozmiarze około 25x25px do rozmiaru wejściowej warstwy (około 250x250px)



Wnioski

Udało się zbudować system różnicowania typów nowotworów osiągający wyniki zbliżone do **90% F1** zarówno dla metod teksturalnych jak i wykorzystujących uczenie głębokie. Głównymi przeszkodami w budowie systemu okazały się **niezbalansowanie zbioru, mała gęstość informacji** (ograniczona ilość informacji) niesionej przez **małe, szarościowe obrazy o niskim kontraście**.

Bardzo duży wpływ na końcowy wynik ma **sposób generacji ROI** z pełnej ramki badania CT. Te same metody generacji cech potrafią dawać wyniki odległe od siebie aż o 15pp przy zmianie metodu generacji.

Niestety gorsze wyniki osiągane były w zadaniu klasyfikacji, gdy badania pacjentów były niezależnie rozdzielone między zbiorami trenującymi i testującymi. **Duża wariancja osobnicza** przypadków nowotworów w połączeniu z **ograniczonym zbiorem przykładów** nie pozwala na zbudowanie tak zgeneralizowanego systemu.



Dalsze badania

Temat różnicowania i przetwarzania obrazów tomograficznych kontynuuję jako temat moich badań doktoranckich. Zajmuję się następującymi tematami:

- nowe metody generacji ROI
- zespoły teksturalno-neuronowe
- augmentacja zbiorów poprzez interpolację oraz metody generatywne
- interpolacyjne metody zmiany rozdzielczości, super-resolution, deblurring i ich wpływ na tekstury i gradienty
- wyjaśnialność decyzji modeli głębokich metodami typu Grad-CAM
- sieci 3D-CNN
- problem segmentacji obrazów tomograficznych



Podziękowania

Chciałabym serdecznie podziękować mojemu promotorowi, prof. dr hab. inż. Tomaszowi Markiewiczowi za inspirację, pomoc i dobre rady w trakcie przeprowadzania badań jak również za możliwość korzystania z jego doświadczenia w dziedzinie zastosowań algorytmów sztucznej inteligencji w medycynie.

Chciałabym również podziękować Narodowemu Centrum Nauki, dzięki którego finansowaniu te badania mogły zostać przeprowadzone.

Dziękuję mojemu tacie, prof. dr hab. inż. Stanisławowi Osowskemu za zachęcenie mnie do poszerzania swoich horyzontów w dziedzinie uczenia maszynowego oraz towarzyszenia mi w pierwszych krokach w tej dziedzinie.

Dziękuję mojemu mężowi i całej rodzinie za wsparcie i motywację do wybranej pracy.

Bibliografia I

- [1] B. Wojtyniak and P. Goryński. *Sytuacja zdrowotna ludności Polski i jej uwarunkowania*. Warszawa, 2018.
- [2] Cancer Research UK - Kidney cancer, 2020 (dostęp 1 Czerwca, 2020).
- [3] Holger Moch, Antonio L. Cubilla, Peter A. Humphrey, Victor E. Reuter, and Thomas M. Ulbright. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs—Part A: Renal, Penile, and Testicular Tumours. *European Urology*, 70(1):93–105, July 2016.
- [4] Bansari Shah, Charmi Sawla, Shraddha Bhanushali, and Poonam Bhogale. Kidney Tumor Segmentation and Classification on Abdominal CT Scans. *International Journal of Computer Applications*, 164(9):1–5, April 2017.
- [5] L. Mredhula and M. A. Dorairangaswamy. Detection and Classification of Tumors in CT Images, 2015.
- [6] Marius George Linguraru, Shijun Wang, Furhawn Shah, Rabindra Gautam, James Peterson, W. Marston Linehan, and Ronald M. Summers. Automated noninvasive classification of renal cancer on multiphase CT: Automated noninvasive classification of renal cancer. *Medical Physics*, 38(10):5738–5746, September 2011.
- [7] Burak Kocak, Aytul Hande Yardimci, Ceyda Turan Bektas, Mehmet Hamza Turkcanoglu, Cagri Erdim, and Ugur Yucetas. Textural differences between renal cell carcinoma subtypes: Machine learning-based quantitative computed tomography texture analysis with independent external validation. *European Journal of Radiology*, 107:149–157, October 2018.

Bibliografia II

- [8] Siva P. Raman, Yifei Chen, James L. Schroeder, Peng Huang, and Elliot K. Fishman. CT Texture Analysis of Renal Masses. *Academic Radiology*, 21(12):1587–1596, December 2014.
- [9] Huseyin Polat and Homay Danaei Mehr. Classification of Pulmonary CT Images by Using Hybrid 3D-Deep Convolutional Neural Network Architecture. *Applied Sciences*, 9(5):940, March 2019.
- [10] Wafaa Alakwaa, Mohammad Nassef, and Amr Badr. Lung Cancer Detection and Classification with 3D Convolutional Neural Network (3D-CNN). *International Journal of Advanced Computer Science and Applications*, 8(8), 2017.
- [11] Wenqing Sun, Bin Zheng, and Wei Qian. Computer aided lung cancer diagnosis with deep learning algorithms. page 97850Z, San Diego, California, United States, March 2016.
- [12] P. Shanmugavadiu and V. Sivakumar. Fractal Dimension Based Texture Analysis of Digital Images. *Procedia Engineering*, 38:2981–2986, 2012.
- [13] Robert M. Haralick, K. Shanmugam, and Its'Hak Dinstein. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621, November 1973.
- [14] Alceu Ferraz Costa, Gabriel Humpire-Mamani, and Agma Juci Machado Traina. An Efficient Algorithm for Fractal Analysis of Textures. In *2012 25th SIBGRAPI Conference on Graphics, Patterns and Images*, pages 39–46, August 2012. ISSN: 2377-5416.

Bibliografia III

- [15] Michael Unser. Local linear transforms for texture measurements. *Signal Processing*, 11(1):61–79, July 1986.
- [16] Michael Unser. Sum and Difference Histograms for Texture Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(1):118–125, January 1986.
- [17] T. Ojala, M. Pietikainen, and D. Harwood. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In *Proceedings of 12th International Conference on Pattern Recognition*, volume 1, pages 582–585, Jerusalem, Israel, 1994. IEEE Comput. Soc. Press.
- [18] N. Dalal and B. Triggs. Histograms of Oriented Gradients for Human Detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 886–893, San Diego, CA, USA, 2005. IEEE.
- [19] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, September 1995.
- [20] Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. *Introduction to data mining*. Pearson Addison Wesley, Boston, 1st ed edition, 2006.
- [21] Y. Bengio. Learning Deep Architectures for AI. *Foundations and Trends® in Machine Learning*, 2(1):1–127, 2009.
- [22] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015.

Bibliografia IV

- [23] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [24] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, May 2017.
- [25] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, Las Vegas, NV, USA, June 2016. IEEE.
- [26] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–9, Boston, MA, USA, June 2015. IEEE.
- [27] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the Inception Architecture for Computer Vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2818–2826, Las Vegas, NV, USA, June 2016. IEEE.
- [28] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alex Alemi. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. *arXiv:1602.07261 [cs]*, August 2016. arXiv: 1602.07261.