

Convolution Neural Network (CNN) Untuk Pengklasifikasian Citra Makanan Tradisional

Akhmad Rohim¹, Yuita Arum Sari², Tibyani³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya
Email: ¹akhmad_rohim@student.ub.ac.id, ²yuita@ub.ac.id, ³tibyani@ub.ac.id

Abstrak

Masyarakat dalam era digital masa kini memfoto sebelum makan merupakan salah satu budaya gaya hidup. Kemudian hasil foto yang didapat akan diunggah ke media sosial. Penyebaran foto makanan tradisional yang masih kurang dalam mengidentifikasi mendorong penelitian ini untuk melakukan penelitian tentang klasifikasi citra makanan tradisional. Pada ekstraksi fitur klasifikasi citra makanan merupakan hal yang sulit karena makanan bisa secara dramatis bervariasi dalam penampilan seperti bentuk, tekstur, warna, dan sifat visual lainnya. Convolution Neural Network (CNN) merupakan metode yang dapat mempelajari sendiri fitur pada citra yang kompleks. Diharapkan hasil evaluasi CNN untuk pengklasifikasian citra makanan tradisional dapat memberikan solusi untuk mengidentifikasi citra makanan tradisional. Hasil penelitian ini menunjukkan dalam membangun arsitektur model Convolutional Neural Network untuk pengklasifikasian citra makanan tradisional membutuhkan 4 layer konvolusional, 4 layer maxpooling dan 2 layer Fully connected. Arsitektur tersebut didapatkan karena mendapatkan nilai *loss value* terkecil dengan nilai 0.000044 pada epoch ke 15 saat proses pembelajaran dan mendapatkan nilai 73% presisi, 69% recall dan 69% Fscore.

Kata kunci: *Convolutional Neural Network, CNN, Makanan, Makanan Tradisional, Confusion Matrix.*

Abstract

People in this digital era take a picture before eating is one of lifestyle. Then the result of the picture will be uploaded to social media. Traditional food's pictures dissemination still less identified encourages this research about the classification of traditional food's image. Extraction of classification features food image is difficult because of food can vary dramatically in appearances such as shape, texture, color, and other visual properties. Convolution Neural Network (CNN) is a method that can learn its own features on a complex image. Hopefully, CNN evaluation results for the classification image of traditional food can provide a solution to identify the image of traditional food. Result of this research in building the architecture of the Convolutional Neural Network model for classification of the traditional food image required 4 conditional layers, 4 max-pooling layers, and 2 fully connected layers. That architecture obtained because it gets the smallest loss value with 0.000044 value on the 15 epoch during the learning process and gets a 73% precision, 69% recall, and 69% F-score.

Keywords: *Convolutional Neural Network, CNN, Food, Traditional Food, Confusion Matrix.*

1. PENDAHULUAN

Makanan mencerminkan karakteristik suatu suku bangsa (Ferguson, 2010). Makanan bisa merepresentasikan suatu negara, dan menjadi simbol penyebaran budayanya. Berbagai negara mengupayakan penyebaran kulinernya secara masif dan ekstensif ke

berbagai negara lainnya. Makanan khas atau makanan tradisional telah dianggap sebagai duta promosi dan bentuk pendekatan kultural yang efektif. Banyak orang Indonesia yang belum pernah ke Jepang, tapi sudah terbiasa mengkonsumsi sushi dan ramen. Termasuk Italia, banyak orang Indonesia yang gemar menyantap pizza atau pasta.

Di masyarakat yang sudah mapan, kebutuhan untuk memuaskan rasa lapar adalah sebuah gaya hidup (Feist & Handriatno, 2010). Memfoto sebelum Makan merupakan salah satu budaya gaya hidup di era digital masa kini. Kemudian hasil foto yang didapat akan diunggah ke media sosial. Dengan adanya budaya seperti itu banyak foto makanan tradisional yang tersebar di media sosial, namun terkadang kita lupa bahkan tidak mengetahui nama makanan tersebut. Maka penyebaran memperkenalkan makanan tradisional akan sedikit terhambat.

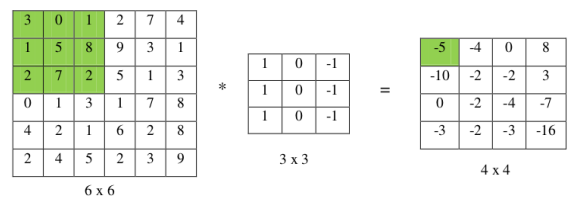
Penyebaran foto makanan tradisional yang masih kurang dalam mengidentifikasi mendorong penelitian ini untuk melakukan penelitian tentang klasifikasi citra makanan tradisional. klasifikasi dalam penelitian ini menggolongkan citra makanan untuk mengidentifikasi makanan yang ingin diketahui, berdasarkan ciri khas setiap citra.

Pada ekstraksi fitur klasifikasi citra makanan merupakan hal yang sangat sulit karena makanan bisa secara dramatis bervariasi dalam penampilan seperti bentuk, tekstur, warna, dan sifat visual lainnya (He, Xu, Khanna, Boushey, & Delp, 2013). Berbagai metode klasifikasi punya cara masing-masing dalam mengekstraksi fitur. (Kumari & Saxena, 2018) mengulas beberapa hasil penelitian mengenai segmentasi dan klasifikasi tumor otak. Dalam ulasannya Kumari mendeskripsikan kelebihan setiap penelitian dan memberikan komentar untuk penelitian lanjut, bahkan Kumari dapat menyimpulkan kelebihan dan kekurangan setiap metode yang digunakan untuk deteksi tumor otak. *Convolution Neural Network* (CNN) merupakan metode yang dapat mempelajari sendiri fitur pada citra yang kompleks (Kumari & Saxena, 2018).

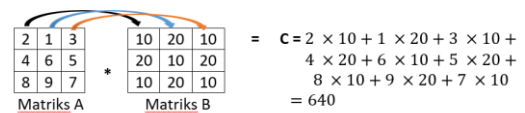
2. LANDASAN KEPUSTAKAAN

2.1 Konvolusi

Konvolusi (convolution) adalah sebuah proses dimana citra dimanipulasi dengan menggunakan eksternal mask atau subwindows untuk menghasilkan citra yang baru. Secara matematis menurut (Madenda, 2015) konvolusi adalah jumlah total dari hasil kali antara setiap elemen yang bersesuaian (memiliki posisi koordinat yang sama) dalam dua matriks atau dua vektor, seperti yang ditunjukkan Gambar 1.



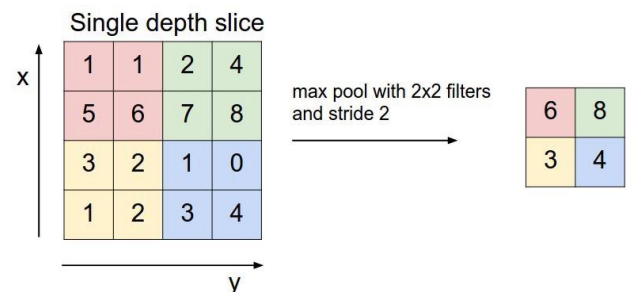
Gambar 1. Konvolusi



Gambar 2 Perhitungan Kovolusi

2.2 Pooling

Pooling adalah proses mereduksi ukuran sebuah data citra. Dalam pengolahan citra, pooling juga bertujuan untuk meningkatkan invariansi posisi dari fitur serta mempercepat komputasi dan mengontrol terjadinya overfitting. Proses pooling dilakukan seperti Gambar 3.



Gambar 3. Pooling
Sumber : (Pokharna, 2016)

2.3 ReLu Aktivasi

Rectified linear unit, yang dikenal sebagai ReLU, adalah cara paling umum dan dasar untuk memperkenalkan non-linearitas ke dalam jaringan saraf. Fungsi ini hanya $\max(0, x)$.

2.4 Confusion Matrix

Confusion matrix adalah salah satu metode yang digunakan untuk mengevaluasi metode-metode klasifikasi. Tabel 1 merupakan gambaran sederhana untuk mempermudah pemahaman tentang istilah confusion matrix dalam keluaran klasifikasi.

Tabel 1. Istilah *Confusion Matrix*
 Sumber : (Sokolova & Lapalme, 2009)

		Kelas Prediksi (<i>Observation</i>)	
		Positif	Negatif
Kelas Aktual (<i>Expection</i>)	Positif	TP	FN
	Negatif	FP	TN

Nilai True Negative (TN) adalah data yang di klasifikasi dengan tepat sebagai keluaran negatif atau salah. True Positive (TP) adalah data yang diklasifikasi dengan tepat sebagai keluaran positif atau benar. False Positive (FP) adalah data yang diklasifikasi dengan kurang tepat apabila keluaran berupa positif atau benar. False Negative (FN) adalah data yang diklasifikasi dengan kurang tepat.

$$\text{Precision} = \frac{\sum_i^n \frac{TP_i}{TP_i + FP_i}}{n} \tag{1}$$

Persamaan (1) merupakan perhitungan rata-rata nilai *precision* yaitu dari data hasil klasifikasi seberapa banyak data yang benar. antara nilai sebenarnya dengan prediksi yang diberikan oleh sistem.

$$\text{Recall} = \frac{\sum_i^n \frac{TP_i}{TP_i + FN_i}}{n} \tag{2}$$

Persamaan (2) merupakan perhitungan rata-rata nilai *recall* yaitu dari seluruh data benar seberapa banyak data yang keluar dalam hasil klasifikasi. Evaluasi *recall* digunakan apabila lebih memilih nilai *False Positif* daripada *False Negatif* (Ghoneim, 2019). Seperti pada memprediksikan seseorang terkena HIV atau tidak, bahwa memprediksikan seseorang yang sehat berlabel HIV lebih dipilih daripada melabelkan sehat kepada seseorang yang terkena HIV.

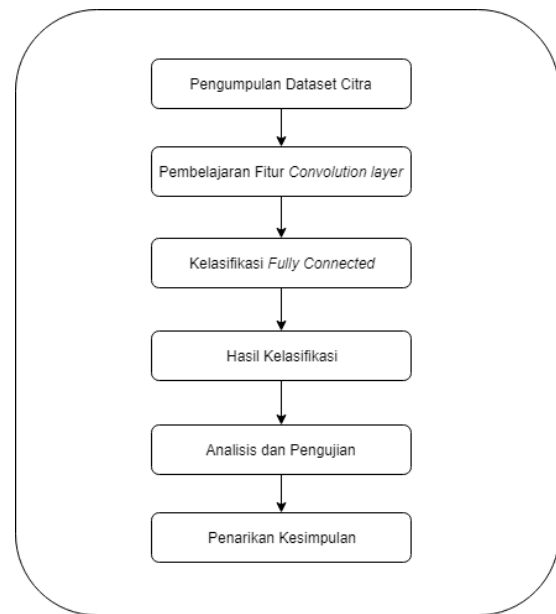
$$\text{Akurasi} = \frac{\sum_i^n \frac{TP_i + TN_i}{TP_i + FP_i + TN_i + FN_i}}{n} \tag{3}$$

$$\text{Fscore} = \frac{\sum_i^n \frac{\text{recall}_i \times \text{presisi}_i}{\beta(\text{recall}_i) + (1-\beta)(\text{presisi}_i)}}{n} \tag{4}$$

Persamaan (3) merupakan perhitungan rata-rata nilai akurasi untuk menunjukkan tingkat efektifitas per kelas dari sebuah klasifikasi (Sokolova & Lapalme, 2009). Sedangkan persamaan (4) merupakan perhitungan rata-rata nilai *Fscore* yang

merupakan nilai kombinasi dari perhitungan recall dan presisi. Akurasi merupakan acuan bagus apabila data keluaran simetris. Seperti contoh berikut apabila TP = 10, FP = 9990, FN = 0, TN = 0. Maka presisinya 0.001, recallnya 1.0, dan akurasinya 0.1. Akurasinya rendah sedangkan semua data sebenarnya telah terprediksi. Jadi penggunaan *Fscore* dibutuhkan dari pada akurasi pada permasalahan tersebut.

3. METODOLOGI PENELITIAN



Gambar 4. Diagram Alur

Berdasarkan Gambar 4 Alur pengerjaan penelitian ini dimulai dari pengumpulan data sampel yang digunakan sebagai input untuk data training, validasi dan pengujian/tes. Kemudian merancang jaringan dengan metode CNN untuk melakukan klasifikasi pada citra makanan. Rancangan jaringan CNN diaplikasikan dengan data training, agar komputer dapat belajar mengenali objek. Untuk mengetahui pembelajaran jaringan diperoleh hasil yang baik dalam membedakan jenis objek makanan, maka jaringan tersebut kemudian dilakukan uji coba terhadap data validasi. Apabila data validasi juga menunjukkan hasil yang baik maka jaringan dapat digunakan untuk klasifikasi pada data tes.

4. PENGUMPULAN DATA

Data yang digunakan dalam penelitian ini adalah data primer yang berupa citra yang diambil menggunakan teknik sebagai berikut.

- a. Citra berwarna berekstensi *.jpg
- b. Citra diambil langsung menggunakan

- kamera smartphone Meizu M4.
- c. Citra diambil citra makanan tradisional sebanyak 1 buah, 2 buah, dan saat makanan terpotong, dengan teknis pengambilan citra seperti berikut.
- d. Jenis citra makanan terdiri dari 20 jenis makanan tradisional yaitu Wingko babat, Gethuk, Roti Bakar, Lumpia, Bakso Bakar, Lemper, Bugis, Onde-onde, Klepon, Tetel, Ketela Goreng, Kue Kucur, Putu Ayu, Molen, Bikang, Kue Sus, Pukis, Serabi, Pisang Ijo, dan Risoles.
- e. Data citra dibagi menjadi dua bagian 280 citra untuk data latih dan 100 citra untuk data uji
- f. Gambar 5 dan Gambar 6 merupakan contoh pengambilan citra



Gambar 5. Contoh pengambilan citra 1



Gambar 6. Contoh pengambilan citra 2

5. HASIL DAN PEMBAHASAN

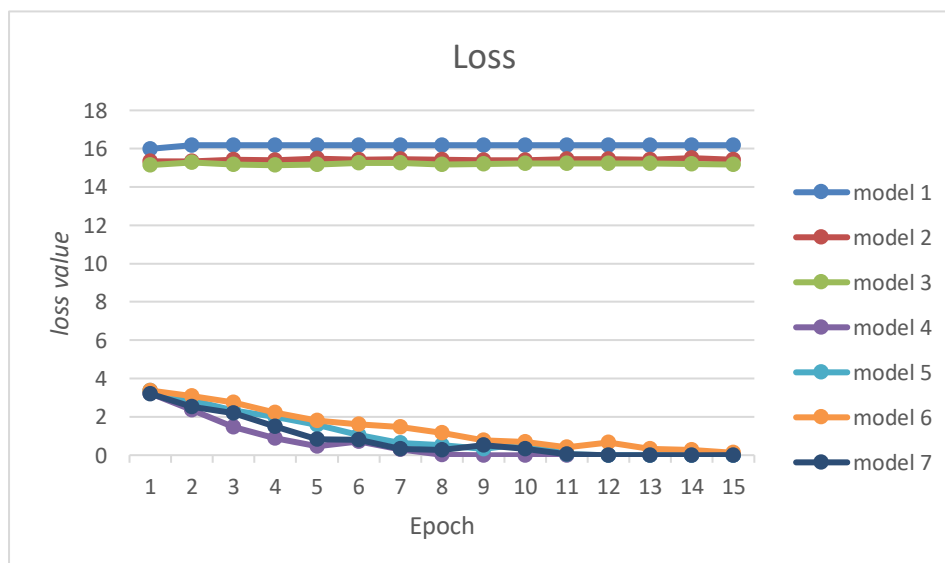
5.1 Pengujian Loss

Pengujian *loss* dan akurasi bertujuan untuk mengetahui seberapa cepat sistem mendapatkan bobot terbaik. Selain itu, pengujian dilakukan untuk mengetahui pengaruh *layer* dalam mencari bobot terbaik. Model yang digunakan dalam pengujian ini ditunjukkan Tabel 2.

Tabel 2. Model CNN

Nama Model	Convolution Layer	Maxpooling Layer	Fully connected Layer
Model 1	1	1	2
Model 2	2	2	2
Model 3	3	3	2
Model 4	4	4	2
Model 5	5	5	2
Model 6	6	6	2
Model 7	7	7	2

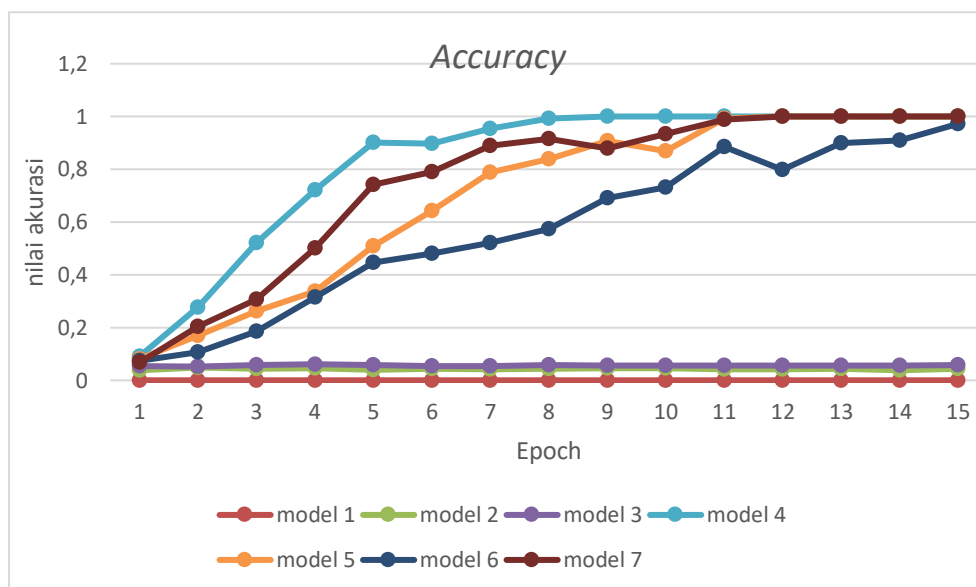
Berdasarkan Tabel 2 *Convolution Layer* dan *Maxpooling layer* jumlahnya sama agar mudah dipelajari polanya. Dalam prosesnya *maxpooling* selalu setelah *convolution layer*. Sedangkan *epoch* yang digunakan untuk melakukan pembelajaran sebanyak 15 kali. Gambar 7 menunjukkan hasil perkaman *loss* value dari proses pembelajaran sistem.



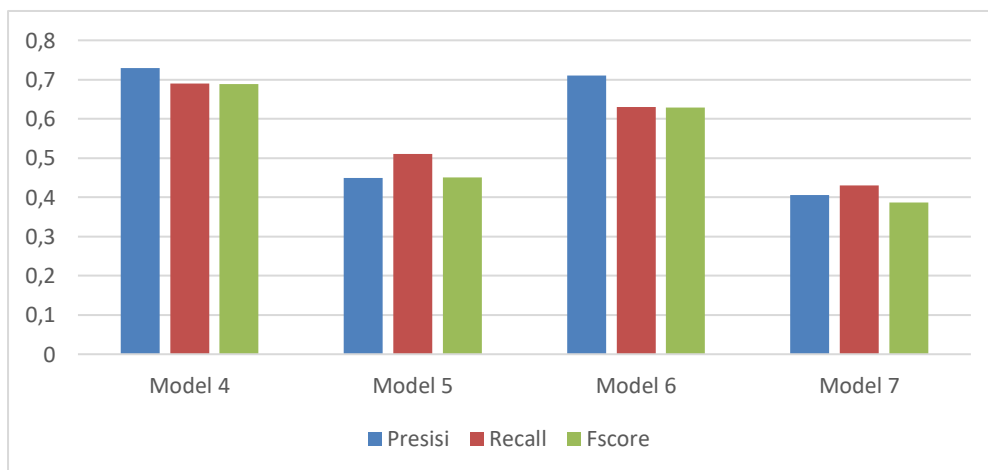
Gambar 7. Grafik Perekaman Loss

Berdasarkan Gambar 7 dapat dilihat bahwa penggunaan layer konvolusi kurang dari 4 layer akan mengalami *loss* yang besar dan konvergen setiap epochnya yaitu sebesar 15,00 sampai 16,50. Model 1 memiliki *loss* selalu besar karena hanya terdapat 1 layer konvolusi didalam arsitekturnya, memiliki layer yang terlalu sedikit membuat sistem akan sulit untuk menemukan fitur khas pada setiap citra. Selain memiliki layer konvolusi yang sedikit, Model 1 juga memiliki layer pooling hanya 1, yang mengakibatkan data terlalu kompleks sehingga akan terjadinya underfitting. Maka penggunaan CNN dengan Model 1, Model 2, dan Model 3 tidak dapat mempelajari gambar apapun. Jadi Model 4, Model 5, Model 6, dan Model 7 dapat memahami data latih dengan baik dibuktikan dengan nilai akurasi pada Gambar 8.

Gambar 8 menyatakan bahwa hampir semua model yang memiliki *layer* konvolusi dan *pooling* lebih dari 3 memiliki akurasi sebesar 1. Memiliki akurasi sebesar 1 dapat dikatakan bahwa sistem mampu memisahkan seluruh data latih berdasarkan kelasnya masing-masing. Karena fungsi dari akurasi adalah validasi proses data latih. *Layer* konvolusi dan *pooling* yang dibutuhkan yaitu minimal 4 agar sistem dapat mempelajari citra. Sedangkan semakin besar epoch maka nilai *loss* dan akurasi akan semakin baik akan tetapi membuat sistem berproses semakin lama. Model 4 lebih cepat dalam memahami fitur citra dibuktikan dengan nilai akurasi yang paling besar setiap *epoch*. Maka berdasarkan evaluasi nilai *loss* dan akurasi dengan 280 data latih yang terdiri dari 20 kelas, arsitektur terbaik menggunakan 4 *layer* konvolusi, 4 *layer* pooling dan 2 *layer* fully connected.



Gambar 8. Grafik Perakaman Akurasi



Gambar 9. Grafik Hasil Rata-rata Pengujian

5.2 Confusion Matrix

Gambar 9 menerangkan bahwa model 4 merupakan model yang terbaik dari model yang lainnya dengan nilai *Fscore* rata-ratanya 68.8%. Model 6 memiliki nilai yang mendekati Model 4 dengan nilai *Fscore* rata-ratanya 62.8%. Permasalahan yang terjadi disemua model adalah dalam pengambilan data latih hanya terdiri dari beberapa sudut pandang yang dilakukan berulang dan tidak sama setiap kelasnya. Pengambilan dengan sedikit sudut pandang akan membuat sistem kurang bisa dalam mempelajari ciri-ciri dari makanan, dan pengambilan secara berulang dalam satu sudut pandang akan membuat *overfitting* dini. Selain pengambilan data, proses pelatihan terlalu lama akan meningkatkan nilai *error* saat pengujian, yang dapat menurunkan nilai *recall* dan presisi.

6. KESIMPULAN

Berdasarkan hasil pengujian dan analisis yang sudah dilakukan, didapatkan kesimpulan Dalam membangun asitektur model *Convolutional Neural Network* untuk pengklasifikasian citra makanan tradisional membutuhkan 4 *layer* konvolusional, 4 *layer maxpooling* dan 2 *layer Fully connected*. Arsitektur tersebut didapatkan karena mendapatkan nilai *loss value* terkecil dengan nilai 0.000044 pada epoch ke 15 saat proses pembelajaran dan mendapatkan nilai 73% presisi, 69% *recall* dan 69% *Fscore*.

Dalam penelitian ini, citra makanan yang memiliki objek berbeda sudut pandang akan memiliki hasil yang berbeda pula. Diharapkan untuk penelitian selanjutnya data set diperbanyak dengan cara citra makanan di foto

dari berbagai sudut. Banyaknya sudut pandang dimungkinkan untuk meningkatkan nilai presisi dan *recall*.

7. DAFTAR PUSTAKA

- Feist, J., & Handriatno. (2010). *teori kepribadian*. Jakarta: Salemba Humanika.
- Ferguson, P. P. (2010). *Culinary Nationalism. Gastronomica the Journal of Food and Culture*, 102-109.
- Ghoneim, S. (2019, April 2). *Accuracy, Recall, Precision, F-Score & Specificity, which to optimize on? Retrieved from Towards Data Science*.
- He, Y., Xu, C., Khanna, N., Boushey, C. J., & Delp, E. J. (2013). *Food Image Analysis. Segmentation, Identification And Weight Estimation*.
- Kumari, N., & Saxena, S. (2018). *Review of Brain Tumor Segmentation and Classification*.
- Madenda, S. (2015). *Pengolahan Citra & Video Digital*. Jakarta: Erlangga.
- Sokolova, M., & Lapalme, G. (2009). *Information Processing and Management. A systematic analysis of performance measures for classification tasks*, 427-437.
- Pokharna, H. (2016, Juli 29). *The best explanation of Convolutional Neural Networks on the Internet! Retrieved from Medium*.