

Multilabel Image Annotation Menggunakan Metode Convolutional Neural Network

Tugas Akhir

diajukan untuk memenuhi salah satu syarat

memperoleh gelar sarjana

dari Program Studi Teknik Informatika

Fakultas Informatika

Universitas Telkom

1301144046

Naufal Dzaky Anwari



Program Studi Sarjana Teknik Informatika

Fakultas Informatika

Universitas Telkom

Bandung

2018

LEMBAR PENGESAHAN

Multilabel Image Annotation Menggunakan Metode Convolutional Neural Network

Multilabel Image Annotation Using Convolutional Neural Network

NIM: 1301144046

Naufal Dzaky Anwari

Tugas akhir ini telah diterima dan disahkan untuk memenuhi sebagian syarat memperoleh

gelar pada Program Studi Sarjana Teknik Informatika

Fakultas Informatika

Universitas Telkom

Bandung, 28 Februari 2018

Menyetujui

Pembimbing I

Pembimbing II

Anditya Arifianto, S.T., M.T.

NIP: 14890028

Jondri, S.Si, M.Si

NIP: 95700035

Ketua Program Studi
Sarjana Teknik Informatika,

Said Al Faraby, S.T., M.Sc.

NIP: 15890019

LEMBAR PERNYATAAN

Dengan ini saya, Naufal Dzaky Anwari, menyatakan sesungguhnya bahwa Tugas Akhir saya dengan judul ” *Multilabel Image Annotation Menggunakan Metode Convolutional Neural Network*” beserta dengan seluruh isinya adalah merupakan hasil karya sendiri, dan saya tidak melakukan penjiplakan yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan. Saya siap menanggung resiko/sanksi yang diberikan jika dikemudian hari ditemukan pelanggaran terhadap etika keilmuan dalam buku TA atau jika ada klaim dari pihak lain terhadap keaslian karya.

Bandung, 28 Februari 2018

Yang Menyatakan,

Naufal Dzaky Anwari

Multilabel Image Annotation Menggunakan Metode Convolutional Neural Network

Naufal Dzaky Anwari¹, Anditya Arifianto,S.T.,M.T.², Jondri,S.Si, M.Si³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹naufalyai@student.telkomuniversity.ac.id, ²anditya@telkomuniversity.ac.id,

³jondri@telkomuniversity.ac.id

Abstrak

Dengan berkembangnya sosial media terutama yang memiliki fitur untuk mengunggah foto dan gambar menyebabkan banyaknya gambar yang diunggah pada sosial media. Banyak gambar yang diunggah tidak diberikan label atau *tag* sesuai dengan isi dari citra yang diunggah, sehingga sangat sulit untuk dikelola, yang dapat digunakan untuk pencarian gambar berbasis isi gambar. Untuk dapat mewujudkan sistem pencarian gambar berbasis isi maka setiap obyek pada gambar harus dikenali terlebih dahulu. Jika pengenalan obyek tersebut dilakukan secara manual maka akan sangat sulit karena akan memakan waktu yang lama dan makna dari setiap orang terhadap suatu gambar berbeda, yang menimbulkan pengenalan obyek yang subyektif. Oleh karena itu dibangunlah sistem penganotasian gambar secara otomatis. Dalam penelitian ini diajukan sebuah metode *Convolutional Neural Network* untuk menangani sistem penganotasian gambar multilabel. metode *Convolutional Neural Network* telah terbukti memiliki performansi yang baik pada kasus klasifikasi gambar, ditunjukkan dengan performansi pada kasus klasifikasi gambar pada ILSVRC yang semakin membaik setiap tahunnya. Performansi tertinggi terhadap data uji pada penelitian ini adalah 81.24%.

Kata kunci : *Convolutional Neural Network, image annotation, multilabel*

Abstract

With the development of social media, especially that has a feature to upload photos and images cause many of images uploaded on social media. Many of the uploaded images are not labeled or tagged in accordance with the content of the uploaded image, so it is very difficult to manage, which can then be used for content-based image search. To be able to realize the content-based image search system then each object in the image must be recognized first. If the annotation of the object is done manually it will be very difficult because it will take a long time and meaning of each person toward an image is different so it will cause the recognition of images is subjective. Therefore, automatic image annotation system was built. In this research, we proposed a Convolutional Neural Network method to handle multilabel image annotation system. *Convolutional Neural Network* method has been shown to have a good performance in the case of image classification, it is shown with performance in classification cases in ILSVRC which is better every years. The highest performance toward test data in this research is 81.24%.

Keywords: *Convolutional Neural Network, image annotation, multilabel*

1. Pendahuluan

Dengan berkembangnya media sosial terutama media sosial yang memiliki fitur untuk mengunggah foto dan gambar seperti Instagram, Facebook, dan Flickr menyebabkan banyak gambar diunggah oleh pengguna sosial media tersebut. Dari gambar-gambar pengguna media sosial yang diunggah, kebanyakan tidak dianotasi oleh pengguna sehingga sangat sulit untuk dikelola. Untuk dapat mewujudkan sistem pencarian gambar berbasis isi maka setiap obyek pada gambar harus dikenali terlebih dahulu dan disimpan. Jika pengenalan obyek tersebut dilakukan secara manual maka akan sangat sulit karena akan memakan waktu yang lama dan makna dari setiap orang terhadap suatu gambar berbeda sehingga menimbulkan pengenalan obyek yang subyektif. Oleh sebab itu dikembangkanlah pengenalan dan penganotasian gambar secara otomatis yang diharapkan dapat mengenali obyek pada gambar secara obyektif berdasarkan ciri visual yang ada dalam gambar.

Pada kasus *Automatic Image Annotation*, banyak penelitian yang telah dilakukan hanya berfokus pada *single-label* saja dimana satu citra/gambar diberikan hanya 1 label. Namun, pada penerapan dalam dunia nyata, satu gambar dapat dikaitkan dengan beberapa *tag* atau label sehingga tidak dapat dilakukan penganotasian hanya dengan *single-label* saja.

Dalam penelitian ini diajukan sebuah metode yang bernama *Convolutional Neural Network* untuk menangani sistem pengantorian gambar yang memiliki multilabel. *Convolutional Neural Network* merupakan tipe jaringan syaraf tiruan yang menggunakan struktur tertentu seperti, *convolutional layer*, *pooling layer*, dan *fully connected layer*. Metode ini telah menunjukkan hasil yang menjanjikan pada permasalahan tentang klasifikasi gambar[2].

Beberapa batas yang terdapat pada penelitian ini adalah :

1. dataset yang digunakan adalah dataset gambar NUS-WIDE-SCENE yang merupakan dataset gambar yang telah diberi label dengan ukuran gambar 224x224 piksel,
2. total dataset yang digunakan sejumlah 27,535 citra yang dibagi menjadi data latih sebanyak 10,913 gambar dan data uji sebanyak 16,622,
3. jumlah kelas sebanyak 30 kelas.

Penelitian bertujuan untuk membangun sistem pengantorian gambar multilabel menggunakan metode *Convolutional Neural Network* dengan mengetahui parameter-parameter yang meningkatkan performa sistem. Selain itu penelitian ini akan membandingkan pengeluaran jumlah label pada sistem dan penggunaan *threshold* pada penentuan label pada sistem. Performa sistem akan diukur dengan menggunakan akurasi sistem.

Urutan penulisan laporan ini adalah sebagai berikut : bagian 2 menunjukkan penelitian terkait dengan penelitian ini. Sistem yang diajukan untuk *Multilabel Image Annotation* menggunakan *Convolutional Neural Network* akan dijelaskan pada bagian 3. Pada bagian 4 akan didiskusikan mengenai hasil pengujian dan evaluasi sistem. Akhirnya, kesimpulan akan dipaparkan pada bagian 5.

2. Studi Terkait

2.1 Image Annotation

Image annotation atau anotasi gambar merupakan cabang dari *image retrieval* yang digunakan untuk memberikan label atau *tag* pada gambar dengan sekumpulan kata kunci berdasarkan isi dari gambar [8]. *Image annotation* menghasilkan label-label yang dapat digunakan untuk pengelompokan gambar berdasarkan isi dari gambar tersebut agar mudah dikelola. Pada penelitian sebelumnya, sistem pengantorian gambar dikembangkan menggunakan berbagai metode seperti metode ekstraksi ciri *Speed-Up Robust Features*(SURF) dengan menggunakan *classifier Support Vector Machine*, menggunakan menggunakan metode klasifikasi *k-Nearest Neighbor(k-NN)* dikombinasikan dengan *Multi Non-negative Matrix Factorization* dan *Semantic Co-occurrence* untuk meningkatkan performa dari sistem yang dibangun [8, 9].

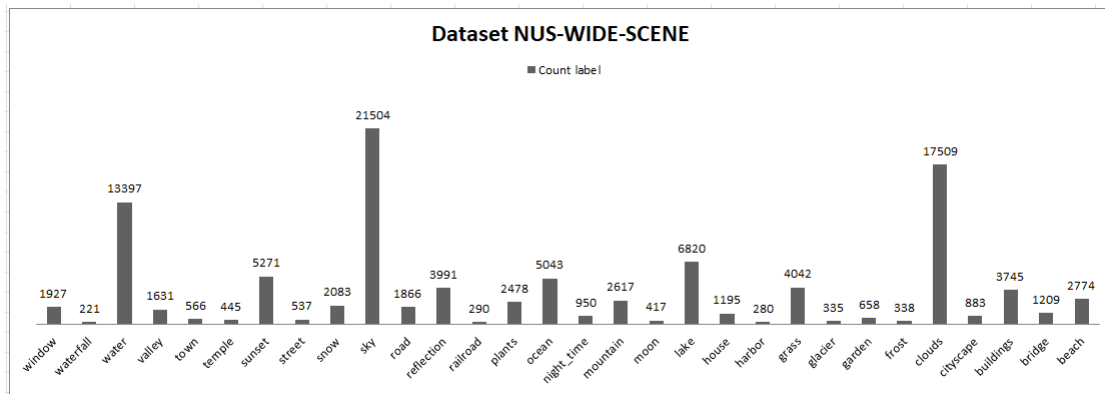
2.2 Convolutional Neural Network

Convolutional Neural Network adalah salah satu dari metode *Deep Learning*. Metode ini sangat efektif digunakan pada pengaplikasian *computer vision* namun tidak menutup kemungkinan bahwa metode ini juga dapat digunakan untuk menyelesaikan kasus *pattern recognition* dan *natural language processing*. Desain dari *Convolutional Neural Network* terinspirasi oleh mekanisme visual pada otak. Terdapat 3 bagian dalam *Convolutional Neural Network* yaitu *convolutional layer*, *pooling layer*, dan *fully-connected layer*. Ketiga lapisan tersebut memiliki fungsi yang berbeda [2].

Beberapa penelitian mengenai *Convolutional Neural Network* telah dilakukan dan berbagai arsitektur *Convolutional Neural Network* telah diperkenalkan. Mulai dari arsitektur AlexNet yang diperkenalkan oleh Alex Krizhevsky, Ilya Sutskever, dan Geoffrey Hinton pada tahun 2012 yang digunakan untuk menyelesaikan kasus klasifikasi gambar pada kompetisi ILSVRC 2012 [4]. Setelah kemunculan AlexNet, penelitian mengenai *Convolutional Neural Network* dan arsitekturnya mulai dikembangkan kembali dalam kasus klasifikasi gambar dengan munculnya berbagai arsitektur seperti VGG pada tahun 2014 dan GoogleNet pada tahun 2015 [5, 7].

2.3 NUS-WIDE Dataset

NUS-WIDE dataset merupakan dataset gambar yang dibuat oleh Laboratorium *Media Search* NUS. Dataset ini terdiri dari 269,648 gambar dan *tag* yang terkait yang diambil dari *Flickr*. Dataset NUS-WIDE dibagi kembali menjadi beberapa kategori agar dapat digunakan untuk analisis visual untuk *image annotation* dan *image retrieval* dikarenakan dataset yang terlalu besar. Salah satu dari kategori tersebut adalah NUS-WIDE-SCENE yang terdiri dari 34.926 gambar dengan *tag* atau label terkait[1]. Gambar 1 merupakan total perhitungan label dari Dataset NUS-WIDE-SCENE dengan 30 kelas yang akan digunakan untuk penelitian *Multilabel Image Annotation* menggunakan *Convolutional Neural Network*.

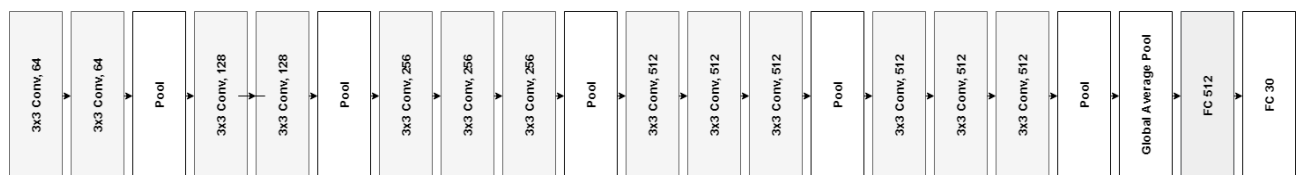


Gambar 1. Total label pada dataset NUS-WIDE-SCENE

3. Sistem yang Dibangun

3.1 Arsitektur Sistem

Arsitektur jaringan *Convolutional Neural Network* yang dibangun menggunakan arsitektur VGGNet 16 namun terdapat beberapa perubahan pada arsitektur VGG16 yang dibangun.

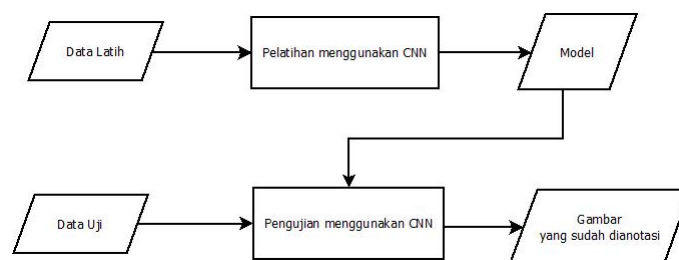


Gambar 2. Arsitektur jaringan *Convolutional Neural Network* pada sistem yang dibangun

Pada gambar 2 arsitektur VGG16 yang telah ada, *layer* klasifikasi pada arsitektur tersebut dihilangkan dan diganti dengan *layer Global Average Pooling* dan *fully-connected layer* dengan jumlah *neuron* 512 dan pada akhirnya ditambah dengan *fully-connected layer* dengan jumlah *neuron* sesuai jumlah kelas. Fungsi aktivasi pada *layer* terakhir diubah menjadi *Sigmoid* untuk kebutuhan klasifikasi multilabel. Pada penelitian ini digunakan *batch normalization* pada *fully-connected layer* sebelum non-linearitas yang bertujuan untuk menyamakan distribusi pada input setiap *layer* sehingga memungkinkan penggunaan *learning rate* yang tinggi dan mempercepat proses pelatihan[3]. Pada arsitektur sistem diterapkan juga *dropout* yang berguna untuk mencegah *overfit* pada jaringan *Convolutional Neural Network*[6].

3.2 Gambaran Umum Sistem

Pada penelitian ini dataset dibagi menjadi 2 yang digunakan untuk proses pelatihan sistem dan pengujian sistem. Gambar diagram 3 menunjukkan gambaran umum dari sistem, dimulai dari pelatihan model menggunakan *Convolutional Neural Network* hingga mendapatkan model dan diuji performansinya pada tahap pengujian.





Gambar 3. Sistem yang dibangun

Dataset pelatihan NUS-WIDE-SCENE yang digunakan adalah dataset dengan label minimal 3 dengan total data pelatihan sebesar 10913 gambar. Gambar yang memiliki label kurang dari 3 tidak akan digunakan pada

penelitian ini. Selanjutnya, data latih yang telah disiapkan dilatih menggunakan jaringan pada *layer* klasifikasi (dilakukan *transfer learning*) selama 100 *epoch* dengan *learning rate* sebesar 0.001 dan menggunakan parameter *update adam optimizer* terhadap 30 kelas. tabel 1 merupakan contoh citra masukan dan hasil anotasi dari sistem.

Tabel 1. Dataset dan label hasil anotasi

Citra	Label
	sunset, sky, clouds
	water, reflection, lake

3.3 Pelatihan menggunakan *Convolutional Neural Network*

Pada penelitian ini dilakukan 2 metode pelatihan yaitu melakukan pelatihan dari awal dan melakukan *transfer learning*

3.3.1 Pelatihan tanpa *transfer learning*

Pada pelatihan tanpa *transfer learning*, arsitektur jaringan yang telah dibangun sesuai gambar 2, dilatih terhadap data latih sebanyak 10913 dengan ukuran *batch* sebesar 5, *learning rate* sebesar 0.01 dan dilatih menggunakan *batch normalization* dan tidak menggunakan *batch normalization* dengan parameter *update* yaitu *Adam Optimizer* selama 60 *epoch*.

3.3.2 Pelatihan menggunakan *transfer learning*

Proses *transfer learning* ini bertujuan untuk menemukan model *layer* klasifikasi yang sesuai dengan data-set latih dengan menggunakan hasil *learning* dari model lain yang telah dilakukan pelatihan. *Transfer learning* dilakukan menggunakan model yang telah dilatih menggunakan data citra ImageNet yang dihilangkan *layer* klasifikasinya dan disambung dengan *layer* klasifikasi yang disesuaikan dengan penelitian ini yaitu *global average pooling*, *fully-connected layer* sebanyak 512 *neuron* dan pada *layer* terakhir dipasang *fully-connected layer* dengan *neuron* sebanyak jumlah kelas yaitu 30 *neuron*. Sistem dilatih menggunakan *batch* sebesar 15, dengan *dropout* 0, 0.50, dan 0.75 serta menggunakan *learning rate* 0.001 dengan *Adam optimizer* selama 100 *epoch*.

3.3.3 Proses Klasifikasi

Pada proses klasifikasi, dilakukan 3 cara penentuan label keluaran yaitu dengan mengeluarkan 3 label terbaik, 5 label terbaik, dan *threshold* pada masing-masing kelas. *threshold* bertujuan untuk menentukan batas penentuan label pada masing-masing *neuron output*. Penentuan *threshold* terbaik dengan cara menelusuri nilai akurasi pada setiap label dengan percobaan *threshold* antara 0.1 hingga 0.9 dengan penambahan 0.01 pada setiap iterasinya. Sedangkan pada penentuan label 3 dan 5 terbaik dilihat dari hasil keluaran *layer* klasifikasi dan diurutkan berdasarkan 3 atau 5 nilai keluaran *neuron* kelas terbesar.

3.4 Pengukuran Performansi Sistem

Dalam penelitian ini, pengukuran performansi dilakukan untuk mengevaluasi sistem yang telah dibangun. Perhitungan performansi sistem menggunakan akurasi yang dapat dituliskan sebagai berikut :

$$Accuracy = \frac{\sum_{i=1}^m N_i^c}{\sum_{i=1}^m N_i^{p\cap g}},$$

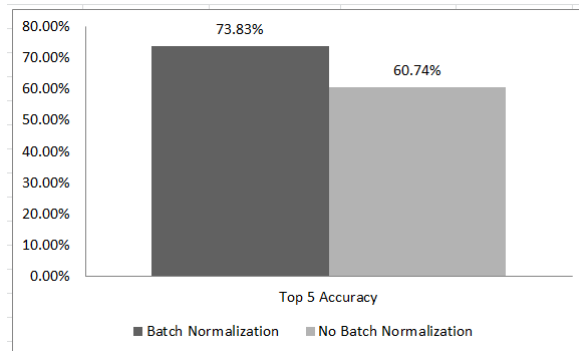
dimana N^c adalah jumlah prediksi label yang benar, $N^{p\cap g}$ merupakan jumlah label yang diprediksi oleh sistem yang beririsan dengan jumlah label sebenarnya. Hasil evaluasi performansi pada setiap skenario pengujian digunakan untuk menentukan parameter terbaik pada sistem yang dibangun.

4. Evaluasi

Pada bagian ini ditampilkan hasil dari pengujian yang telah dilakukan terkait beberapa parameter yaitu *batch normalization*, *dropout*, dan penentuan label top 3, top 5, dan threshold serta ditampilkan pula pelatihan dengan metode *transfer learning*.

4.1 Perbandingan pengujian menggunakan model pelatihan menggunakan *batch normalization* dengan tidak menggunakan *batch normalization*

Dari arsitektur *Convolutional Neural Network* pada gambar 2, diberikan *batch normalization* pada *convolutional layer* dan pada *fully-connected layer* dan dilatih selama 60 *epoch* menggunakan *learning rate* 0.01 dengan *adam optimizer*. Berikut merupakan hasil dari pengujian terhadap arsitektur yang menggunakan *batch normalization* dan tidak menggunakan *batch normalization*.



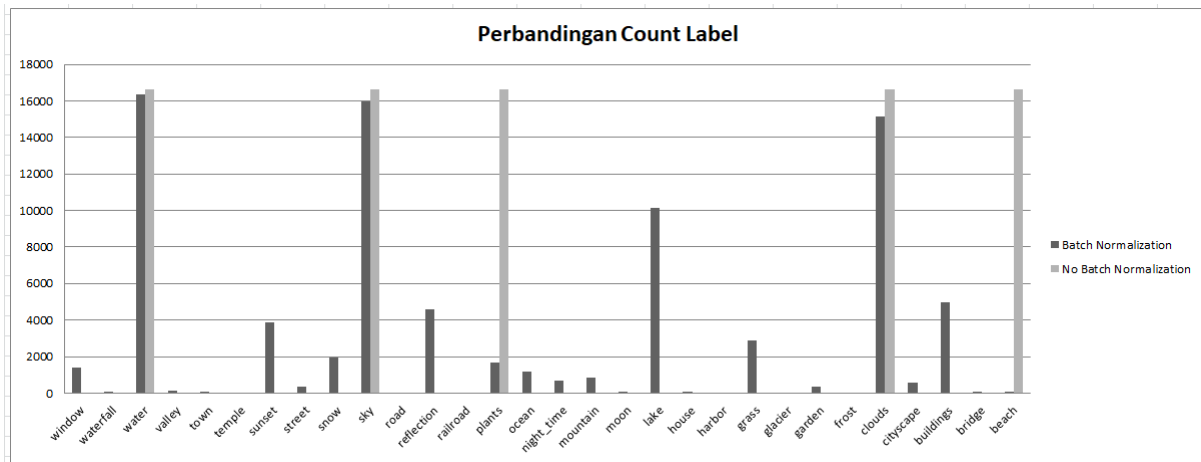
Gambar 4. Perbandingan akurasi top 5 penggunaan *batch normalization*

Dari grafik 4 pada jaringan tanpa *batch normalization* menunjukkan akurasi sebesar 60.74% dengan total gambar yang benar-benar salah dianotasi adalah 1461 gambar. Sedangkan pada jaringan dengan *batch normalization* menunjukkan akurasi sebesar 73.83% dengan total gambar yang benar-benar salah dianotasi sebesar 1018 gambar. Penghitungan frekuensi label pada setiap gambar dilakukan untuk memperjelas mengenai perbedaan pada penggunaan *batch normalization*. Perhitungan tersebut bisa dilihat pada grafik 5.

dari grafik 5 jika dibandingkan dengan grafik dataset 1 menunjukkan bahwa jaringan tanpa *batch normalization* hanya dapat menganotasi 5 label yaitu *water*, *sky*, *plants*, *clouds*, *beach* dan proses pembelajaran dengan 60 *epoch* masih belum dapat mengetahui ciri-ciri yang dimiliki pada label-label yang lain. Sedangkan pada jaringan yang menggunakan *batch normalization* sudah mulai terdapat label-label yang dapat dikenali oleh *Convolutional Neural Network*. Hal ini menunjukkan bahwa penggunaan *batch normalization* pada *convolutional layer* dapat mempercepat *Convolutional Neural Network* dalam proses pembelajaran dikarenakan pada masing-masing *convolutional layer* dilakukan normalisasi sehingga distribusi dari seluruh masukan sebelum masuk ke fungsi aktivasi akan disesuaikan yang membuat perbedaan rentang pada setiap masukan ke fungsi aktivasi tidak akan berbeda jauh.

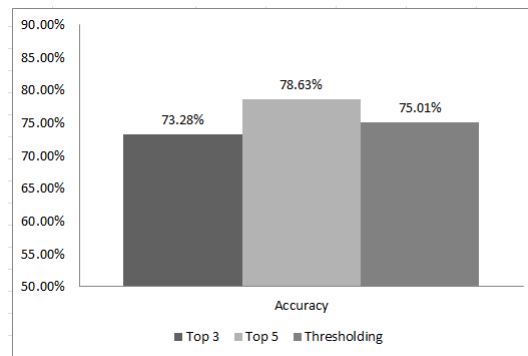
4.2 Pengujian terhadap model yang dilatih menggunakan *transfer learning*

Pada pengujian dengan hasil pelatihan menggunakan *transfer learning*, dilakukan pengujian terhadap mekanisme pengeluaran label yaitu top 3, top 5, dan *threshold*, terhadap jaringan yang dilatih menggunakan *batch*



Gambar 5. Perbandingan frekuensi label *batch normalization*

normalization pada *fully-connected layer* dan tidak, serta penggunaan *dropout* 0, 0.5, dan 0.75 pada jaringan. Hasil pengujian terhadap mekanisme pengeluaran label top 3, top 5, dan *threshold* dapat dilihat pada grafik 6. Pada grafik 6 mekanisme pengeluaran label top 5 memiliki nilai akurasi sebesar 78.63% yang lebih tinggi dari



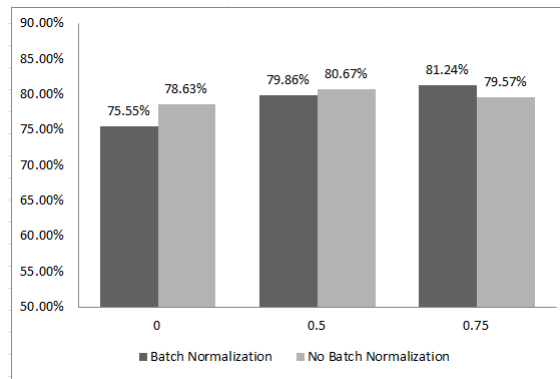
Gambar 6. Perbandingan akurasi pada mekanisme pengeluaran label

top 3 yang memiliki nilai akurasi sebesar 73.28% dan *threshold* sebesar 75.01%. Top 5 memiliki nilai tertinggi dikarenakan pada setiap data citra yang dianotasi, sistem akan selalu mengambil 5 dengan *score* terbaik. Hal ini dikarenakan pada dataset NUS-WIDE-SCENE terdapat data yang memiliki label yang tidak lengkap. Contohnya adalah pada tabel 2 dimana label target hanya 1 yaitu label *road*, sedangkan pada citra tersebut seharusnya dapat memiliki label *building*, *street*, dan *sky*.

Tabel 2. Contoh dataset yang memiliki label kurang lengkap

Citra	Label Target	Label Prediksi
	<i>road</i>	<i>street, sky, road, clouds, buildings</i>

Selanjutnya dilakukan pengujian terhadap penggunaan *dropout* dan *batch normalization* pada *fully-connected layer* pada proses pelatihan *transfer learning*. Berikut merupakan grafik perbandingan penggunaan probabilitas *dropout* 0, 0.5, dan 0.75 dengan *batch normalization* dan tidak menggunakan *batch normalization* terhadap akurasi top 5. Pada grafik 7 terlihat bahwa pada hasil pelatihan *transfer learning*, penggunaan *batch normalization* pada *fully-connected layer* memiliki akurasi lebih rendah terhadap jaringan tanpa *batch normalization* namun



Gambar 7. Perbandingan akurasi top 5 pada penggunaan *dropout* dan *batch normalization*

jika dibandingkan terhadap penambahan probabilitas *dropout*, jaringan dengan *batch normalization* pada probabilitas *dropout* 0 memiliki akurasi top 5 sebesar 75.55%, pada probabilitas *dropout* 0.5 memiliki akurasi top 5 sebesar 79.86%, dan pada probabilitas *dropout* 0.75 akurasi top 5 yang dihasilkan sebesar 81.24% hal ini menunjukkan bahwa pada saat penambahan probabilitas *dropout*, akurasi top 5 jaringan ini selalu naik. Sedangkan pada jaringan tanpa *dropout* akurasi top 5 akan turun pada probabilitas *dropout* 0.75 walaupun penurunan tersebut tidak signifikan yaitu sebesar 1.1%. Hal ini terjadi karena pada saat probabilitas *dropout* 0.75 banyak *neuron* pada *fully-connected layer* akan mati sehingga pada proses pembelajaran distribusi masukan menuju fungsi aktivasi akan berubah karena pada setiap proses pembelajaran banyak *neuron* yang dimatikan yang membuat hasil bobot-bobot *transfer learning* yang masuk pada *fully-connected layer* mengalami perubahan distribusi. Namun dengan dilakukannya *batch normalization* pada probabilitas *dropout* yang tinggi yaitu 0.75 dapat meningkatkan performansi sistem karena *batch normalization* mengatasi perbedaan distribusi pada masing-masing *neuron* pada *fully-connected layer*.

5. Kesimpulan

Pada penelitian *multilabel image annotation*, *Convolutional Neural Network* memiliki performansi yang baik untuk melakukan anotasi gambar dengan akurasi top 5 mencapai 81.24%. *Batch normalization* dapat meningkatkan performansi sistem dengan cara menyesuaikan distribusi pada setiap *neuron* masukan sebelum menuju fungsi aktivasi. Pada penggunaan *batch normalization* pada *convolutional layer*, performansi sistem meningkat dari 60.74% menjadi 73.83%. Mekanisme penentuan label dapat mempengaruhi performansi *Convolutional Neural Network*. Pada pelatihan menggunakan *transfer learning*, pengeluaran label 3 terbaik atau top 3 memiliki akurasi paling kecil sebesar 73.28% dibanding dengan mekanisme *threshold* sebesar 75.01% dan 5 label terbaik atau top 5 nilai akurasi sebesar 78.63%. Mekanisme penentuan label top 5 memiliki akurasi yang paling tinggi dikarenakan mengambil 5 kemungkinan label terbaik. Penggunaan *batch normalization* dan *dropout* pada jaringan *Convolutional Neural Network* dapat meningkatkan performansi sistem hingga 81.24%. Melihat performansi *Convolutional Neural Network* pada sistem pengantasian gambar multilabel yang cukup baik, diharapkan pada penelitian selanjutnya pengembangan sistem anotasi dengan label kelas lebih banyak dengan parameter lain seperti tebal dan ukuran *filter* serta arsitektur yang berbeda sangat dianjurkan untuk penelitian selanjutnya.

Daftar Pustaka

- [1] T.-S. Chua, J. Tang, R. Hong, H. Li, Z. Luo, and Y.-T. Zheng. Nus-wide: A real-world web image database from national university of singapore. In *Proc. of ACM Conf. on Image and Video Retrieval (CIVR'09)*, Santorini, Greece., July 8-10, 2009.
- [2] Y. Guo et al. Deep learning for visual understanding: review. 2015.
- [3] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *CoRR*, abs/1502.03167, 2015.
- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.

- [5] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.
- [6] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15:1929–1958, 2014.
- [7] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. *CoRR*, abs/1409.4842, 2014.
- [8] N. M. Tuhin Shukla and S. Sharma. Automatic image annotation using surf feature. *International Journal of Computer Applications (0975 - 8887)*, 68:4, 2013.
- [9] F. Zhong and L. Ma. Image annotation using multi-view non-negative matrix factorization and semantic co-occurrence. *Region 10 Conference (TENCON), 2016 IEEE*, 2016.

Lampiran

5.1 Data Observasi

Tabel 3. *Threshold* dengan parameter probabilitas *dropout* 0 dan tanpa *batch normalization*

tag	threshold
window	0.81
waterfall	0.73
water	0.4
valley	0.86
town	0.8
temple	0.89
sunset	0.84
street	0.63
snow	0.6
sky	0.67
road	0.85
reflection	0.8
railroad	0.53
plants	0.3
ocean	0.72
night_time	0.61
mountain	0.89
moon	0.72
lake	0.24
house	0.83
harbor	0.81
grass	0.45
glacier	0.88
garden	0.87
frost	0.83
clouds	0.3
cityscape	0.48
buildings	0.32
bridge	0.79
beach	0.88

Tabel 4. Percobaan *transfer learning* dengan parameter probabilitas *dropout* dan *batch normalization*

Learning Rate	Epoch	BN	Dropout	Top 3 Accuracy	Top 5 Accuracy	Threshold Accuracy
0.001	100	Yes	0	71.40%	75.55%	73.29%
0.001	100	Yes	0.5	75.59%	79.86%	76.26%
0.001	100	Yes	0.75	76.75%	81.24%	77.49%
0.001	100	No	0	73.28%	78.63%	75.01%
0.001	100	No	0.5	76.08%	80.67%	76.68%
0.001	100	No	0.75	75.30%	79.57%	77.56%