

Home > Archives > Vol 12, No 2

Vol 12, No 2

June 2023

DOI: <http://doi.org/10.11591/ijai.v12.i2>

Table of Contents

| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|
| Artificial Intelligence: the major role it played in the management of healthcare during COVID-19 pandemic Tabrez Uz Zaman, Elaf Khalid Alharbi, Aeshah Saleem Bawazeer, Ghala Abdullah Algethami, Leen Abdullah Almeahdi, Taif Muhammed Alshareef, Yasmin Awwadh Alotaibi, Hosham Mohammed Osman Karar | PDF 505-513 |
| Automated invoice data extraction using image processing Akanksh Aparna Marjunathi, Manjunath Sudhakar Nayak, Santhanam Nishith, Satish Nitin Pandit, Shreyas Sunkad, Pratiba Deenadhayalan, Shobha Gangadhara | PDF 514-521 |
| A review of factors that impact the design of a glove based wearable devices Soly Mathew Biju, Obada Al Khatib, Hashir Zahid Sheikh | PDF 522-531 |
| Stability of classification performance on an adaptive neuro fuzzy inference system for disease complication prediction Sri Kusumadewi, Linda Rosita, Elyza Gustri Wahyuni | PDF 532-542 |
| An adjustment degree of fitting on fuzzy linear regression model toward manufacturing income Nurfawahida Ramly, Mohd Saifulah Rusiman, Muhammad Ammar Shafi, Suparman S, Firdaus Mohamad Hamzah, Ozlem Gुरुnlu Alma | PDF 543-551 |
| Masking preprocessing in transfer learning for damage building detection Hapnes Toba, Hendra Bunyamin, Juan Elisha Widayana, Christian Wibisono, Lucky Surya Haryadi | PDF 552-559 |
| Neural network-based pH and coagulation adjustment system in water treatment Oscar Ivan Vargas Mora, Dalam Camilo Parrado Nieto, Jairo David Cuero Ortega, Javier Eduardo Martinez Baquero, Robinson Jimenez Moreno | PDF 560-567 |
| An application of Vietnamese handwriting text recognition for information extraction from high school admission form Pham The Bao, Le Tran Anh Dang, Nguyen Duy Tam, Nguyen Nhat Truong, Pham Cung Le Thien Vu, Trinh Tan Dat | PDF 568-576 |
| An image-based convolutional neural network system for road defects detection Mohamed Anis Benallal, Mustapha Si Tayeb | PDF 577-584 |
| Deep learning speech recognition for residential assistant robot Robinson Jiménez-Moreno, Ricardo A. Castillo | PDF 585-592 |
| Information system based on multi-value classification of fully connected neural network for construction management Tetyana Honcharenko, Roman Alexetrad, Andrii Shipakov, Olexsandr Khomenko | PDF 593-601 |
| Classification of dances using AlexNet, ResNet18 and SqueezeNet1.0 Khalif Amir Zakry, Irwandi Hipiny, Hamimah Ujir | PDF 602-609 |
| Deep convolutional neural networks-based features for Indonesian large vocabulary speech recognition Hilman F. Pardede, Purwoko Adhi, Vicky Zilvan, Ade Ramdan, Dikdik Krisnandi | PDF 610-617 |
| Multi-channel of electroencephalogram signal in multivariable brain-computer interface Esmeralda Contessa Djamal, Dimas Andhika Sury | PDF 618-626 |
| A hybrid approach for face recognition using a convolutional neural network combined with feature extraction techniques Hicham Benradi, Ahmed Chater, Abdelali Lassar | PDF 627-640 |
| Classification of semantic segmentation using fully convolutional networks based unmanned aerial vehicle application Shouket Abdulrahman Ahmed, Hazry Desa, Abadal-Salam T, Hussain | PDF 641-647 |
| New approach for detecting multi-point relays in the optimized link state routing protocol using self-organizing map artificial neural network: OLSR-SOM Omar Barki, Zouhair Guermour, Adnane Addaim | PDF 648-655 |
| An improved artificial bee colony with perturbation operators in scout bees' phase for solving vehicle routing problem with time windows Salah Mortada, Yuhanis Yusuf | PDF 656-666 |
| A collaborated genetic with lion optimization algorithms for improving the quality of forwarding in a vehicular ad-hoc network Sami Abduljabbar Rashid, Mustafa Maad Hamdi, Lukman Audah, Mohammed Ahmed Jubair, Mustafa Hamid Hassan, Mohammed Salah Abood, Salama A. Mostafa | PDF 667-677 |
| Multi-objective load balancing in cloud infrastructure through fuzzy based decision making and genetic algorithm based optimization Neema George, Anoop Balakrishnan Kadan, Vinodh P. Vijayan | PDF 678-685 |
| Robustness enhancement study of augmented positive identification controller by a sigmoid function Abbas H. Issa, Sarab A. Mahmood, Abdulrahim T. Humod, Nihad M. Ameen | PDF 686-695 |
| An efficient security analysis of bring your own device Pullaqura Soubhagyalakshmi, Kalli Satyanarayan Reddy | PDF 696-703 |
| Cross-checked screening application for reliable categorisation of familial hypercholesterolaemia: design and development of the prototype Marshima Mohd Rosli, Muthukkaruppan Annamalai, Noor Alicezah Mohd Kasim, Chua Yung-An, Hapizah Mohd Nawawi | PDF 704-713 |
| A high frame-rate of cell-based histogram-oriented gradients human detector architecture implemented in field programmable gate arrays Syifaui Fuada, Trio Adiono, Hans Kasan | PDF 714-730 |
| K-means clustering analysis and multiple linear regression model on household income in Malaysia Gan Pei Yee, Mohd Saifulah Rusiman, Shuhaida Ismail, Suparman Suparman, Firdaus Mohamad Hamzah, Muhammad Ammar Shafi | PDF 731-738 |
| Dialect classification using acoustic and linguistic features in Arabic speech Mohammed Ali Humayun, Hayati Yassin, Pg Emeroyanffion Abas | PDF 739-746 |
| An investigation of wine quality testing using machine learning techniques Sathishkumar Mani, Reshmy Avanavalappil Krishnakutty, Sabana Swaminathan, PrasannaVenkatesan Theerthagiri | PDF 747-754 |
| Query expansion based on modified Concept2vec model using resource description framework knowledge graphs Sarah Dahir, Abderrahim El Qadi | PDF 755-764 |
| Design and implementation of the web (extract, transform, load) process in data warehouse application Seddiq Q. Abd Al-Rahman, Ekram H. Hasan, Ali Naki Sagheer | PDF 765-775 |
| Hypertension prediction using machine learning algorithm among Indonesian adults Rico Niawan, Budi Utomo, Kemal N. Siregar, Kalamullah Ramlil, Besral Besral, Ruddy J. Suhatri, Okky Assetya Pratiwi | PDF 776-784 |

USER

Username

Password

Remember me

[Login](#)

CITATION ANALYSIS

- Dimensions
- Google Scholar
- Scholar Metrics
- ScimagoJR
- Scopus
- Web of Science
- Scifl

QUICK LINKS

- Editorial Boards
- Reviewers
- Author Guidelines
- Online Submission
- Peer Review Process
- Publication Fee
- Abstracting and Indexing
- Publication Ethics
- Visitor Statistics
- Contact Us
- Registration for IJ-AI's Professional Reviewers

JOURNAL CONTENT

Search

Search Scope

[Search](#)

Browse

- By Issue
- By Author
- By Title

| | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|
| Location-aware hybrid microcosmic routing scheme for mobile opportunistic network Shobha R. Bharamagoudar, Shivakumar V. Saboji | PDF 785-793 |
| Innovations in t-way test creation based on a hybrid hill climbing-greedy algorithm Heba Mohammed Fadhil, Mohammed Aboulah, Mohammed Younis | PDF 794-805 |
| Machine learning classifiers for detection of glaucoma Reshma Verma, Lakshmi Shrinivasan, Basvaraj Hiremath | PDF 806-814 |
| Thai Hom Mali rice grading using machine learning and deep learning approaches Akara Thammasittikul, Jitsaeng Petsuwan | PDF 815-822 |
| Effect of word embedding vector dimensionality on sentiment analysis through short and long texts Mohamed Chiny, Marouane Chihab, Abdelkarim Ait Lahcen, Omar Bencharaf, Younes Chihab | PDF 823-830 |
| Predicting students' academic performance using e-learning logs Malak Abdullah, Mameud Al-Ayyoub, Farah Shalwan, Saif Rawashdeh, Rob Abbott | PDF 831-839 |
| Iban related motif classification with adaptive smoothing Silvia Joseph, Irwandi Hipiny, Hamimah Ujir | PDF 840-850 |
| Cuckoo search algorithm for construction site layout planning Melinda Fitriani, Nur Maghfiroh, Anak Agung Ngurah Penira Redi, Janice Ong, Muhamad Rausingy Fikri | PDF 851-860 |
| Architecting a machine learning pipeline for online traffic classification in software defined networking using spark Sama Salam Samaan, Hassan Awheed Jaead | PDF 861-873 |
| BMSF-ML: big mart sales prediction using different machine learning techniques Rao Faizan Ali, Anqad Kuneer, Ahmed Almaghthawi, Amal Alghamdi, Suliman Mohamed Fati, Ebrahim Abdulwasea Abdullah Ghaleb | PDF 874-883 |
| Insights on assessing image processing approaches towards health status of plant leaf using machine learning Harsha Raju, Veena Kalludi Narasimhaiah | PDF 884-891 |
| Hybrid Forex prediction model using multiple regression, simulated annealing, reinforcement learning and technical analysis Hana Jamali, Younes Chihab, Iván García-Magarino, Omar Bencharaf | PDF 892-911 |
| Product defect detection based on convolutional autoencoder and one-class classification Meryem Chaabi, Mohamed Hamlich, Moncef Garouani | PDF 912-920 |
| Facial expression recognition of masked faces using deep learning Boutaina Hicloud, Mohammed El Haj Tiran | PDF 921-930 |
| Deep learning based object detection in nailfold capillary images Suma Kuncha Venkatapathiah, Sethu Selvi Selvan, Pranav Nanda, Manisha Shetty, Vikas Mallikarjuna Swamy, Kushagra Awasthi | PDF 931-942 |
| Modeling of an artificial intelligence based enterprise callbot with natural language processing and machine learning algorithms Imad Aattouri, Hicham Mounici, Mohamed Rida | PDF 943-955 |
| Combating propaganda texts using transfer learning Malak Abdullah, Dia Abujaber, Ahmed Al-Qarqaz, Rob Abbott, Mirsad Hadzikadic | PDF 956-965 |
| Spectrum sensing using 16-QAM and 32-QAM modulation techniques at different signal-to-noise ratio: a performance analysis Neha Chaudhary, Rashima Mahajan | PDF 966-973 |
| Using skeleton model to recognize human gait gender Omar Ibrahim Alsafi, Saba Qasim Hasan, Abdulfata Hussain Maray | PDF 974-983 |
| Evaluation of massive multiple-input multiple-output communication performance under a proposed improved minimum mean squared error precoding Dheyaa Jasim Kadhim, Muna Hadi Saleh, Sadiq Jassim Abu-Loukh | PDF 984-994 |



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).

IAES International Journal of Artificial Intelligence (IJ-AI)

ISSN/e-ISSN 2089-4872/2252-8938

This journal is published by the [Institute of Advanced Engineering and Science \(IAES\)](#) in collaboration with [Intelektual Pustaka Media Utama \(IPMU\)](#).



[View IJAI Stats](#)

Masking preprocessing in transfer learning for damage building detection

Hapnes Toba¹, Hendra Bunyamin², Juan Elisha Widya¹, Christian Wibisono¹,
Lucky Surya Haryadi¹

¹Master Program in Computer Science, Faculty of Information Technology, Maranatha Christian University, Bandung, Indonesia

²Bachelor Program in Informatics Engineering, Faculty of Information Technology, Maranatha Christian University, Bandung, Indonesia

Article Info

Article history:

Received Jan 12, 2022

Revised Oct 10, 2022

Accepted Nov 9, 2022

Keywords:

Classification

Convolutional neural network

Damage building detection

Image segmentation

Transfer learning

ABSTRACT

The sudden climate change occurring in different places in the world has made disasters more unpredictable than before. In addition, responses are often late due to manual processes that have to be performed by experts. Consequently, major advances in computer vision (CV) have prompted researchers to develop smart models to help these experts. We need a strong image representation model, but at the same time, we also need to prepare for a deep learning environment at a low cost. This research attempts to develop transfer learning models using low-cost masking pre-processing in the experimental building damage (xBD) dataset, a large-scale dataset for advancing building damage assessment. The dataset includes eight types of disasters located in fifteen different countries and spans thousands of square kilometers of satellite images. The models are based on U-Net, i.e., AlexNet, visual geometry group (VGG)-16, and ResNet-34. Our experiments show that ResNet-34 is the best with an F1 score of 71.93%, and an intersection over union (IoU) of 66.72%. The models are built on a resolution of 1,024 pixels and use only first-tier images compared to the state-of-the-art baseline. For future orientations, we believe that the approach we propose could be beneficial to improve the efficiency of deep learning training.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Hapnes Toba

Master Program in Computer Science, Faculty of Information Technology,

Maranatha Christian University

Jl. Suria Sumantri No. 65, Bandung 40164, West Java, Indonesia

Email: hapnestoba@it.maranatha.edu

1. INTRODUCTION

A considerable amount of unprecedented weather changes around the world have made disasters more unpredictable and more severe than before [1]. On the other hand, the advance in machine learning (ML) and computer vision (CV) has brought computer science algorithms the capability of building intelligent and independent solutions for disaster prevention all around the world. Additionally, the increasing availability of satellite images from the United States and European scientific agencies, such as the United States Geological Survey (USGS), National Oceanic and Atmospheric Administration (NOAA), and European Space Agency (ESA) has further cultivated more and more research on ML and CV with the help of domain experts, such as humanitarian assistance and disaster recovery (HADR) and remote sensing experts [2]–[4]. Training accurate and robust CV models needs large-scale and a variety of datasets; moreover, all buildings have different designs from one another. The differences between designs depend on locations or countries where the buildings are located. It may seem a challenge for CV models to recognize all types of building from various places.

The experimental building damage (xBD) dataset [2] comprises satellite images utilized for detecting building shapes and assessing building damages. Furthermore, the dataset encompasses eight types of disasters located in fifteen different countries and covers thousands of square-kilometer satellite images. The dataset consists of pairs of images; specifically, the first and second images represent conditions of a region before and after a disaster respectively. Additionally, the dataset has been annotated in javascript object notation (JSON) form; therefore, there is no need for further annotation processes. This research attempts to build CV models which are capable of detecting and segmenting building shapes on satellite images before and after disasters occur.

One of the important issues in image processing is the complexity during the feature extraction process. In this sense, we need a powerful image representation model, but on the other hand, we also need to prepare for a low-cost deep learning environment. In this research, our main research question is thus, how to prepare a simple yet powerful image preprocessing for transfer learning.

The transfer learning approach has been chosen for the approach of this research because the technique has utilized best practices for state-of-the-art models [5]–[7]. Particularly, the trained models for detecting building shapes from given images employ convolutional neural networks (CNN) architectures such as AlexNet [8], visual geometry group (VGG) [9], and ResNet [10]. Furthermore, we postulate that by using a low complexity pre-processing algorithm, the entire transfer learning process will be more efficient.

2. METHOD

2.1. State-of-the-art techniques

Image segmentation refers to segmenting or partitioning an image into different areas, with each area commonly representing a class. Specifically, CV techniques can be employed on satellite images to extract a partition of the image as an object of a predefined class. Various techniques for satellite image segmentation consist of thresholding, clustering, region-based, and artificial neural networks (ANN). Among those techniques, ANN proves to be giving the best accuracy [11].

CNN is known as one of the deep learning techniques used for CV tasks. Specifically, CNN is developed from multilayer perceptron (MP) to process two-dimensional data such as images [7], [12], [13]. CNN technique has three layers which are divided into two main parts, feature learning, and classifier parts. The feature learning part consists of convolution layers and pooling layers. The classifier part comprises a fully connected layer. Arrangements of CNN shall construct various forms of CNN architectures such as AlexNet [8], VGG [9], and ResNet [10].

U-Net has the capability of processing large-size images and generating outputs whose sizes are the same as the ones of inputs. Another advantage of U-Net is the processing speed which is constant during the training phase. The U-Net training process adopts the CNN training method which replaces a pooling operation with the upsampling operation so the convolutional and pooling layers of the model can return the size of an input image [14]. The u-Net architecture resembles a letter U which is divided into contracting and expansive parts. A contracting part tackles the feature extraction process while an expansive part involves transferring features and reconstructing images to the original input size.

Previous satellite image datasets before xBD only cover one type of natural disaster with various label criteria for damaged buildings [4], [15], [16]. Furthermore, datasets [17], and [18] provide locations of disaster occurrences; however, these datasets do not include damaged building structure images. There are also datasets with multi-view imagery such as change detection and land classification [19]–[21] where several visits to one site and a time series of satellite images are provided. Prominent satellite image segmentation techniques are applied to road segmentation; specifically, the techniques are unsupervised [22], [23]. However, there are limited amounts of literature that discuss road segmentation and identification with obstructions. Other segmentation approaches to detect damaged buildings propose a ML model trained on non-building shapes. [24]. Ronneberger *et al.* [14] develop a U-Net architecture whose model is specifically designed to segment objects in medical images with a limited size of training data. They employ both the Glioblastoma-astrocytoma U373 cells on a polyacrylamide substrate (PhC-U373) and the Henrietta Lacks cells on a flat glass recorded by differential interference contrast microscopy (DIC-HeLa) datasets to measure the model's intersection over union (IoU) value. The IoU values for PhC-U373 and DIC-HeLa datasets are 0.9203 and 0.7756 respectively. Gupta *et al.* [2] establish a baseline model for the xBD dataset. Particularly, they utilize SpaceNet, a variant of U-Net architecture as shown in Figure 1. The IoU values of their model for ground and building are 0.97 and 0.66 respectively. Kurama *et al.* [11] use U-Net architecture trained on 2,000 images of the defence science and technology laboratory (DSTL) dataset and achieve 98% accuracy.

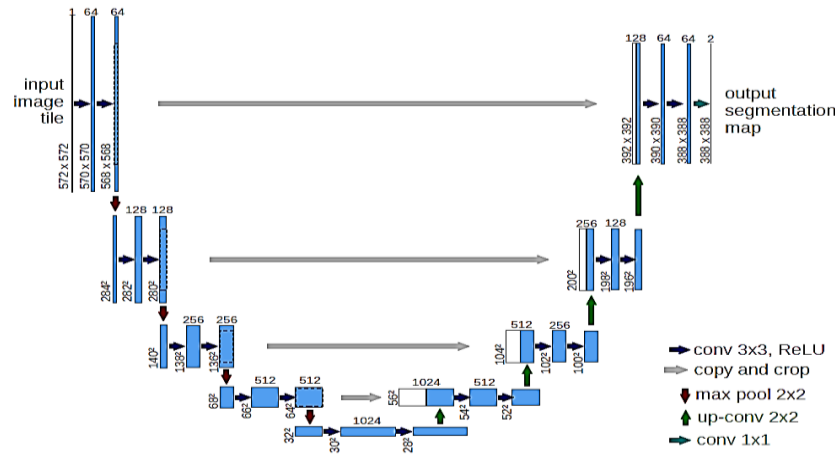


Figure 1. U-Net architecture [10]

2.2. Contributions

This research contributes to CV recent literature in the following aspects:

- We experimented with a lightweight masking preprocessing procedure for the disaster images in the xBD dataset which gives low complexity yet powerful feature extraction in the U-Net architectures.
- We compare several variants of CNN U-Net architectures utilized for detecting building shapes before and after disasters from the xBD dataset. The CNN segmentation techniques analyzed in this research are AlexNet, VGG-16, and ResNet-34 as these techniques are the most widely used in the literature [5].

We believe that this research shall give some insights into the masking preprocessing procedure and its potential during transfer learning. As far as we know. Our research is the first which compares the original experiment in the xBD dataset in various U-Net architectures.

2.3. Experiments

2.3.1. Dataset

This research uses the xBD dataset which is one of the publicly available annotated satellite images with high resolution. The dataset has more than 850,000 polygons for 22,000 building images from six types of disasters worldwide, which encompass more than 45,000 square kilometers [2]. The dataset annotations are done by experts in their fields such as California air national guard (CAL FIRE) and federal emergency management agency (FEMA). Each satellite image has red green blue (RGB) values which form three squares of 1,024 pixels. In this research, the first tier of the dataset is used and divided by xView2 into two portions, train and validation set. The number of images in the train set and validation set is 5,598 and 1,866 respectively which consist of the types of disasters described in Table 1.

Table 1. Number of images for each disaster

| Disaster | Number of images | |
|----------------------|------------------|------------|
| | Train | Validation |
| guatemalare-volcano | 36 | 10 |
| hurricane-orence | 638 | 238 |
| hurricane-harvey | 638 | 190 |
| hurricane-matthew | 476 | 188 |
| hurricane-michael | 686 | 218 |
| mexico-earthquake | 242 | 68 |
| midwest-flooding | 558 | 172 |
| palu-tsunami | 226 | 82 |
| santa-rosa-wildfire | 452 | 154 |
| socal-fire 1,646 546 | 1,646 | 546 |
| Total | 5,598 | 1,866 |

2.3.2. Image preprocessing

The xBD dataset annotations are saved into JSON format and one of the annotations is building information coordinates on an image. Furthermore, this coordinate information is preprocessed into creating

a masking image [25]. The masking image consists of two classes, which are ground and building. A zero-value pixel in a masking image refers to a ground; on the other hand, a one-value pixel indicates a building. Figures 2 and 3 show an image before and after the masking process is applied. Furthermore, the masking image is used as a label or target during the training of a CV model.



Figure 2. An image before masking



Figure 3. An image after masking is applied

2.3.3. Model training

A model (f) is trained on satellite images to detect buildings at pixel levels shows in Algorithm 1, that is:

Algorithm 1 Preprocessing images algorithm

```

1: procedure Preprocessing (images, json_file)
2: read the json_file containing building coordinates
3: for each image in images do
4: for each pixel (i, j) in the image do
5: if (i, j) is part of a building then #utilize the JSON file
6: (i, j) = 1
7: else
8: (i, j) = 0
    
```

For every pixel in an image, p_{ij} with (i, j) as the coordinate of the pixel. This training method is a well-known technique known as image segmentation in CV literature [26]. We opt to choose the transfer learning approach as this approach gives the best performance results which are elaborated by Raffel *et al.* [27]. The convolutional base of CNN has been trained on the ImageNet dataset [5]; therefore, the xBD dataset is normalized by the statistics of ImageNet to have the same range of input distribution [28]. An illustration of the transfer learning approach is Figure 4.

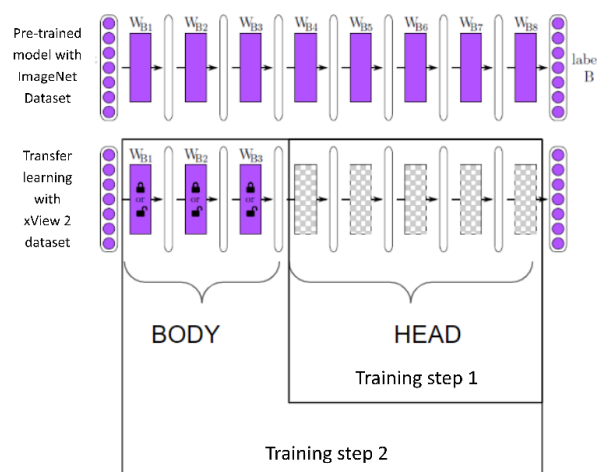


Figure 4. Transfer learning approach illustration

The transfer learning approach utilizes a convolutional base learner which has learned a lot of features from a dataset for a specific task. Next, this knowledge will be used to perform the task on a different dataset without initializing weights randomly. If the dataset is quite large, the weights of the model can be updated wholly; this training process is commonly called fine-tuning. Similarly, our model undergoes a two-stage training process. Firstly, only the head of the model is trained on the dataset. Next, the model is trained for updating the weights of all layers [29].

The deep learning library which was used during the training is fast.ai which is run on n1-highmem-4 and graphics processing unit (GPU) NVidia tesla T4 of google cloud platform for 4 days the learning rate is 0.0003 obtained from the cyclical learning rate finder algorithm [30]. During training, data augmentation techniques such as flipping images horizontally, rotating images, magnifying images, adjusting brightness, contrasting images, and wrapping images are also used. In addition, the performance parameters for this task are precision, recall, and F1, given in (1)-(3), with true positive (TP), false positive (FP), and false negative (FN) carefully assessed.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (3)$$

Additionally, IoU metric in (4), the metric used in Gupta *et al.* [2], is also utilized to evaluate our model.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \quad (4)$$

3. RESULTS AND DISCUSSION

Three CNN-based architectures, *i.e.*: AlexNet, VGG-16, and ResNet-34, are trained on 512 by 512-pixel images with 10 epochs. Our best-performing models are chosen based on the F1 score because of the imbalance between ground and building image instances in our dataset. The comparison of the three models when only the heads are trained is displayed in Table 2.

The best model among the three models, that is ResNet-34 is trained on 512 and 1,024 pixels on the head only with the number of epochs of 40 and a learning rate of 0.0003. Next, all layers are fine-tuned with a learning rate ranging from 0.000001 to 0.0001. Results of the training process are Tables 3 and 4. Both tables display that the models give better F1 scores and IoU results than the ones in Table 2.

Table 2. Comparison of the three models at the tenth epoch

| Model | Accuracy | Precision | Recall | F1 Score |
|-----------|----------|-----------|--------|----------|
| AlexNet | 0.950 | 0.640 | 0.271 | 0.357 |
| VGG-16 | 0.958 | 0.696 | 0.391 | 0.474 |
| ResNet-34 | 0.966 | 0.700 | 0.674 | 0.683 |

Table 3. Training ResNet-34 model at 512 pixels resolution

| Train | Accuracy | Precision | Recall | F1 Score | Mean IoU Building |
|-------------|----------|-----------|--------|----------|-------------------|
| Head | 0.974 | 0.803 | 0.708 | 0.751 | 0.592 |
| Fine-tuning | 0.975 | 0.804 | 0.720 | 0.758 | 0.609 |

Table 4. Training ResNet-34 model at 1,024 pixels resolution

| Train | Accuracy | Precision | Recall | F1 Score | Mean IoU Building |
|-------------|----------|-----------|--------|----------|-------------------|
| Head | 0.978 | 0.789 | 0.681 | 0.719 | 0.667 |
| Fine-tuning | 0.978 | 0.791 | 0.676 | 0.717 | 0.669 |

Figure 5 presents a sample of our ground truth pixel values, while Figure 6 presents the predictions. The performances of the trained model on the validation set are measured by IoU [14], specifically the IoU building. Table 5 (512 pixels) and Table 6 (1,024 pixels) depict the segmentation results and IoU values of the

validation set from ten disasters. Image segmentation of hurricane-matthew gives the least value while the one of guatemala-volcano surprisingly displays a good result considering the size of its dataset which is the least.

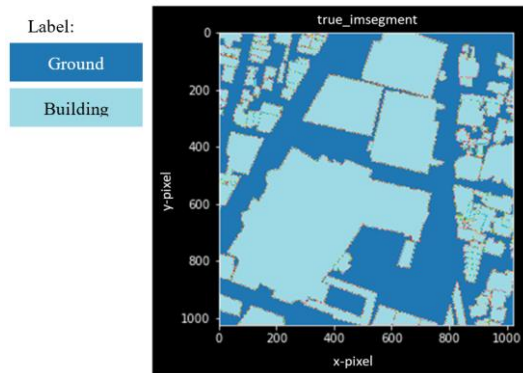


Figure 5. The ground truth pixel values of one sample in the validation set. The image size is 1,024×1,024 pixels (in the x and y-axis directions)

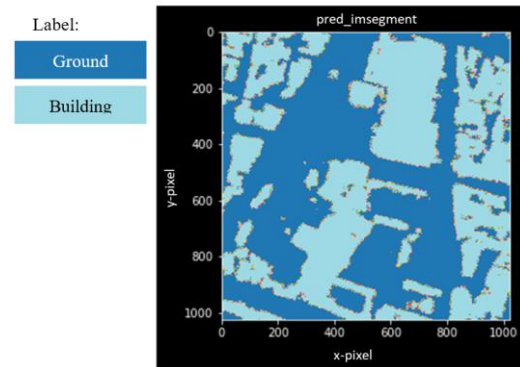


Figure 6. The predicted pixel values of the sample. The image size is 1,024×1,024 pixels (in the x and y-axis directions)

Table 5. IoU of disasters at 512 pixels resolution

| Disaster | IoU segmentation at 512 pixels per disaster | | | |
|---------------------|---------------------------------------------|--------------|-------------|--------------|
| | Training Head | | Fine Tuning | |
| | IoU ground | IoU building | IoU ground | IoU building |
| guatemala-volcano | 0.992716 | 0.516159 | 0.992850 | 0.528130 |
| hurricane-florence | 0.996835 | 0.651713 | 0.996637 | 0.666267 |
| hurricane-harvey | 0.976307 | 0.672333 | 0.975674 | 0.688640 |
| hurricane-matthew | 0.993617 | 0.276589 | 0.993091 | 0.314112 |
| hurricane-michael | 0.986097 | 0.675072 | 0.985711 | 0.689483 |
| mexico-earthquake | 0.905966 | 0.671344 | 0.902866 | 0.687535 |
| midwest- | 0.994258 | 0.640343 | 0.994310 | 0.656130 |
| palu-tsunami | 0.953890 | 0.700680 | 0.947037 | 0.729558 |
| santa-rosa-wildfire | 0.986534 | 0.623966 | 0.986657 | 0.638125 |
| socal-fire | 0.996651 | 0.532794 | 0.996702 | 0.541918 |

Table 6. IoU of disasters at 1,024 pixels resolution

| Disaster | IoU segmentation at 512 pixels per disaster | | | |
|---------------------|---------------------------------------------|--------------|-------------|--------------|
| | Training Head | | Fine Tuning | |
| | IoU ground | IoU building | IoU ground | IoU building |
| guatemala-volcano | 0.995799 | 0.582504 | 0.995598 | 0.577696 |
| hurricane-florence | 0.997853 | 0.744014 | 0.997796 | 0.749505 |
| hurricane-harvey | 0.978948 | 0.734031 | 0.979413 | 0.731891 |
| hurricane-matthew | 0.994308 | 0.364812 | 0.994263 | 0.375385 |
| hurricane-michael | 0.988052 | 0.742830 | 0.987936 | 0.742655 |
| mexico-earthquake | 0.914349 | 0.705674 | 0.916219 | 0.700831 |
| midwest- | 0.996147 | 0.726253 | 0.996176 | 0.726788 |
| palu-tsunami | 0.957746 | 0.742502 | 0.958971 | 0.744839 |
| santa-rosa-wildfire | 0.989383 | 0.708836 | 0.989252 | 0.700055 |
| socal-fire | 0.997107 | 0.611816 | 0.997081 | 0.614974 |

4. CONCLUSION

This research delves into satellite image segmentation using a U-Net architecture with convolutional bases such as AlexNet, VGG-16, and ResNet-34. The final model is ResNet-34 with an accuracy of 0.978409, precision of 0.789098, recall of 0.681466, and F1-score of 0.719300 when the head of the model is trained. The mean of the IoU is 0.667237, and this number is similar to the IoU of our baseline as reported in the initial xBD dataset exploration. However, our research utilizes a smaller dataset, which is only the first tier compared to the baseline. Moreover, our architecture is simpler than the one of the baseline, that is ResNet-34. We also trained the model in 4 days compared to the baseline which is in 7 days. These advantages can be achieved because of the transfer learning approach. For future directions, we believe that our proposed method can be beneficial to improve the training efficiency in deep learning. It is strongly

recommended to cooperate with satellite image experts to obtain in-depth interpretation and information. Furthermore, a greater number of images should also give better performances at detecting buildings from satellite images. Consequently, models can be improved to detect levels of damage to buildings after successful segmentation.

ACKNOWLEDGEMENTS

The research presented in this paper was partially supported by the Research Institute and Community Service (LPPM) at Maranatha Christian University, Indonesia.




REFERENCES

- [1] M. K. Van Aalst, "The impacts of climate change on the risk of natural disasters," *Disasters*, vol. 30, no. 1, pp. 5–18, Mar. 2006, doi: 10.1111/j.1467-9523.2006.00303.x.
- [2] R. Gupta *et al.*, "xBD: A Dataset for Assessing Building Damage from Satellite Imagery," Nov. 2019, [Online]. Available: <http://arxiv.org/abs/1911.09296>.
- [3] S. Dhingra and D. Kumar, "A review of remotely sensed satellite image classification," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 3, p. 1720, Jun. 2019, doi: 10.11591/ijece.v9i3.pp1720-1731.
- [4] R. Foulser-Piggott, R. Spence, R. Eguchi, and A. King, "Using remote sensing for building damage assessment: GEOCAN study and validation for 2011 Christchurch earthquake," *Earthquake Spectra*, vol. 32, no. 1, pp. 611–631, Feb. 2016, doi: 10.1193/051214EQS067M.
- [5] F. Zhuang *et al.*, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021, doi: 10.1109/JPROC.2020.3004555.
- [6] M. S. AL-Huseiny and A. S. Sajit, "Transfer learning with GoogLeNet for detection of lung cancer," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 2, p. 1078, May 2021, doi: 10.11591/ijeecs.v22.i2.pp1078-1086.
- [7] M. Moe Htay, "Feature extraction and classification methods of facial expression: a survey," *Computer Science and Information Technologies*, vol. 2, no. 1, pp. 26–32, Mar. 2021, doi: 10.11591/csit.v2i1.p26-32.
- [8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017, doi: 10.1145/3065386.
- [9] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," vol. 1, Sep. 2014, [Online]. Available: <http://arxiv.org/abs/1409.1556>.
- [10] H. Imaduddin, F. Yusufida Ala, A. Fatmawati, and B. A. Hermansyah, "Comparison of transfer learning method for COVID-19 detection using convolution neural network," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 2, pp. 1091–1099, Apr. 2022, doi: 10.11591/eei.v11i2.3525.
- [11] V. Kurama, S. Alla, and S. Tumula, "Detection of natural features and objects in satellite images by semantic segmentation using neural networks," *Artificial Intelligence Techniques for Satellite Image Analysis*, pp. 161–188, 2020, doi: 10.1007/978-3-030-24178-0_8.
- [12] J. Wu, "Introduction to convolutional neural networks," *National Key Lab for Novel Software Technology*, vol. 5, no. 23, p. 495, 2017.
- [13] Y. H. Liu, "Feature extraction and image recognition with convolutional neural networks," *Journal of Physics: Conference Series*, vol. 1087, p. 62032, Sep. 2018, doi: 10.1088/1742-6596/1087/6/062032.
- [14] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234–241.
- [15] W. Shi, M. Zhang, R. Zhang, S. Chen, and Z. Zhan, "Change detection based on artificial intelligence: state-of-the-Art and challenges," *Remote Sensing*, vol. 12, no. 10, p. 1688, May 2020, doi: 10.3390/rs12101688.
- [16] S. A. Chen, A. Escay, C. Haberland, T. Schneider, V. Staneva, and Y. Choe, "Benchmark dataset for automatic damaged building detection from post-hurricane remotely sensed imagery," Dec. 2018, [Online]. Available: <http://arxiv.org/abs/1812.05581>.
- [17] L. Giglio, J. T. Randerson, and G. R. van der Werf, "Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4)," *Journal of Geophysical Research: Biogeosciences*, vol. 118, no. 1, pp. 317–328, Mar. 2013, doi: 10.1002/jgrg.20042.
- [18] I. Demir *et al.*, "DeepGlobe 2018: a challenge to parse the earth through satellite images," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Jun. 2018, pp. 172–17209, doi: 10.1109/CVPRW.2018.00031.
- [19] J. Ding *et al.*, "Object detection in aerial images: A large-scale benchmark and challenges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 1, 2021, doi: 10.1109/TPAMI.2021.3117983.
- [20] D. Ienco, R. Gaetano, C. Dupaquier, and P. Maurel, "Land cover classification via multitemporal spatial data by deep recurrent neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1685–1689, Oct. 2017, doi: 10.1109/LGRS.2017.2728698.
- [21] D. Peng, Y. Zhang, and H. Guan, "End-to-End change detection for High desolution satellite images using improved UNet++," *Remote Sensing*, vol. 11, no. 11, p. 1382, Jun. 2019, doi: 10.3390/rs11111382.
- [22] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual U-Net," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 749–753, May 2018, doi: 10.1109/LGRS.2018.2802944.
- [23] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, "Albumentations: Fast and flexible image augmentations," *Information (Switzerland)*, vol. 11, no. 2, 2020, doi: 10.3390/info11020125.
- [24] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: a review," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3212–3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.
- [25] H. Bunyamin, "Creating mask images from shapely polygons," *Creating Mask Images from Shapely Polygons*. 2022, Accessed: Jun. 28, 2022. [Online]. Available: https://hbunyamin.github.io/computer-vision/Creating_Masks/.
- [26] J. Guo *et al.*, "GluonCV and gluon NLP: Deep learning in computer vision and natural language processing," *Journal of Machine Learning Research*, vol. 21, no. 23, pp. 1–7, 2020.
- [27] C. Raffel *et al.*, "Exploring the limits of transfer learning with a unified text-to-text transformer," *Journal of Machine Learning Research*, vol. 21, no. 140, pp. 1–67, 2020.




- [28] P. Mettes, D. C. Koelma, and C. G. M. Snoek, "Shuffled imageNet banks for video event detection and search," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 16, no. 2, pp. 1–21, May 2020, doi: 10.1145/3377875.
- [29] J. Howard and S. Gugger, "Fastai: A layered api for deep learning," *Information (Switzerland)*, vol. 11, no. 2, 2020, doi: 10.3390/info11020108.
- [30] L. N. Smith, "No more pesky learning rate guessing games," *CoRR Abs150601186*, vol. 5, p. 363, 2015.

BIOGRAPHIES OF AUTHOR






Hapnes Toba    graduated in 2002 with a Master of Science from the Delft University of Technology in the Netherlands and completed his doctoral degree in computer science at Universitas Indonesia in 2015. He is an associate professor in the area of artificial intelligence and is interested in information retrieval, natural language processing, and computer vision. He has been a faculty member in the Faculty of Information Technology at Maranatha Christian University since 2003. He is also an active board member of the Indonesian Computational Language Association (INACL) and the vice-chair of the Information and Communication Technology Forum (FTIK) of the Association of Christian Universities and Colleges in Indonesia (BK-PTKI). He can be contacted by email at: hapnestoba@it.maranatha.edu.






Hendra Bunyamin    is an assistant professor at the Faculty of Information Technology at Maranatha Christian University, Bandung, Indonesia. He graduated from the Mathematics Department at Bandung Institute of Technology in 1999 and pursue his master's degree from the Software Engineering, Informatics Department at the same university in 2005. He is very passionate about teaching mathematics and programming. His research focuses on the application techniques of automatic learning algorithms. He can be contacted by email at: hendra.bunyamin@it.maranatha.edu.






Juan Elisha Widyya    graduated in 2012 with a Bachelor of Engineering degree and 2021 with a Master of Computer Science degree, both from Maranatha Christian University. He has an interest in pattern recognition, deep learning, computer vision, and time series analysis. Since 2012 he has been working at PT. Yamaha Indonesia Motor Manufacturing as an Area Service Development. He can be contacted at email: 1979006@maranatha.ac.id.



Christian Wibisono    graduated in 2021 with a Master of Computer Science from Maranatha Christian University. He is currently working at a manufacturing company as an IT Assistant Manager. He is a passionate software engineer specializing in React, GraphQL, Next JS, PL/SQL, and SQL. He is passionately studying machine learning and deep learning technology to expand his knowledge in this field, so soon he can implement this knowledge to support the needs of the business. He can be contacted at email: 1979002@maranatha.ac.id.



Lucky Surya Haryadi    graduated in 2021 with a Master of Computer Science from Maranatha Christian University. Lucky's research is focused on computer vision, social networks, and social media analytics. He decided to pursue his master's degree in computer science to improve small businesses with technological advancement. He is the founder of Mierakigai, a company that works in Social Media Advertising and Social Media Analytics. He can be contacted at email: suryaharyadi@gmail.com and 1979001@maranatha.ac.id.