Deep Learning Methods for Virus Identification from Digital Images

Luxin Zhang and Wei Qi Yan

Auckland University of Technology, Auckland 1010 New Zealand

***Abstract*— The use of deep learning methods for virus identification from digital images is a timely research topic. Given an electron microscopy image, virus recognition utilizing deep learning approaches is critical at present, because virus identification by human experts is relatively slow and time-consuming. In this project, our objective is to develop deep learning methods for automatic virus identification from digital images, there are four viral species taken into consideration, namely, SARS, MERS, HIV, and COVID-19. In this work, we firstly examine virus morphological characteristics and propose a novel loss function which aims at virus identification from the given electron micrographs. We take into account of attention mechanism for virus locating and classification from digital images. In order to generate the most reliable estimate of bounding boxes and classification for a virus as visual object, we train and test five deep learning models: R-CNN, Fast R-CNN, Faster R-CNN, YOLO, and SSD, based on our dataset of virus electron microscopy. Additionally, we explicate the evaluation approaches. The conclusion reveals SSD and Faster R-CNN outperform in the virus identification.**

Keywords*— *classification, localization, CNN, virus, electron microscopy images

I. Introduction

Viruses have a variety of shapes, like ball, silk thread, bullet, brick, tadpole, etc. [1][2][3] A rich assortment of viruses have highly contagious capability and show a significant threat to public health. SARS-CoV-2 has seriously jeopardized health and safety of our human beings in 2020 [4][5]. How to automatically and accurately classify viruses especially those extremely infectious viruses so as to save human labor of medics is a necessity at present.

HIV (i.e., human immunodeficiency virus) is a member of the genus lentivirus having similar morphological characteristics with two species: HIV-1 and HIV-2 [20, 21]. SARS (i.e., severe acute respiratory syndrome) is a member of coronavirus family and shares morphological similarities [22,23], appeared in electron micrograph as pleomorphic spherical particles with bulbous surface projections [24]. MERS-CoV (i.e., Middle East respiratory syndrome-related coronavirus) [25] and MERS, also known as camel flu, are a coronavirus species [26]; Severe acute respiratory syndrome corona virus 2 (SARS-COV-2), or COVID-19, given the year of discovery, is a novel severe coronavirus [27]. Four typical proteins are identified as spikes (S), envelope (E), membrane (M), and nucleocapsid (N) [28] in COVID-19.

In this paper, virus detection utilizing deep learning methods is propounded to identify the viruses through using electron microscope images. The applications of deep learning methods have engendered significant contributions [6]. Convolutional neural network (CNN), as one of the deep learning methods, produces reliable results in image analysis, speech recognition and more than others [7-10]. CNNs manifest numerous advanced outcomes in digital image processing, including object recognition, segmentation, image super-resolution, object detection, etc. [11-14]. Therefore, the goal of this project is to explore and exploit virus-oriented computational methods from the viewpoint of deep learning.

In this paper, digital images having four classes of viruses will be taken into consideration: SARS, MERS, HIV, and COVID-19. Additionally, for visually detecting the viruses by using computer vision and digital image processing, we will develop and compare the performance of five deep learning models, namely, R-CNN (i.e., region-based CNN), Fast R-CNN, Faster R-CNN, YOLO, and SSD (i.e., single shot multibox detector), to distinguish viruses from microscope images.

This paper is organized in the follows. We have our literature review in Section II, our methods are depicted in Section III. Our experimental results and resultant analysis are presented in Section IV. Our conclusion will be drawn in Section V.

II. Literatyre Reviuew

Virus detection from digital images is one of the subjects of visual object detection and recognition in computer vision and digital image processing [15]. This paper will fulfil the comparisons of our experimental results related to object detection. The focus of this paper is on four viruses: HIV, SARS, MERS, and COVID-19. The morphological characteristics of these four viruses are manifested as spherical shape and scattered like “white dots” for spike protein projections [16-19].

In order to select the training model, R-CNN network utilizes image regions of interest (ROI) in an agnostic manner by utilizing a selective search [29,30]. ROIs appear in semantic color, shading, texture, morphological characteristics, scale, etc., selective search approaches solve this issue by using hierarchical grouping [31-33]. Fast R-CNN is an improvement of R-CNN achieved with higher computational speed as well as prediction accuracy [34-37]. In order to accommodate Fast R-CNN for virus recognition and locating, numerous amendments were conducted to the internal structure of Fast R-CNN. Faster R-CNN is a further improvement of R-CNN to get faster computation and higher prediction accuracy [38]. A region proposal network (RPN) was proposed in combination with Fast R-CNN and incorporates attention mechanisms [39,40].

YOLO (You Only Look Once), a novel CNN structure, has achieved fast computations as well as distinguished prediction accuracy [41,42]. Numerous versions of YOLO model were updated to gratify more for the field of virus classification and locating in electron microscopy images. A research study [43] aptly demonstrated excellent prediction accuracy for SSD in comparison with others, YOLO was applied to deliver more reliable result with digital images

having minor sizes (i.e., low resolution with reduced visual information) [44-46]. Consequently, the performance of identifying viruses from microscope images was compared with deep learning models like R-CNN, Fast R-CNN, Faster R-CNN, YOLO, and SDD.

III. Our Methods

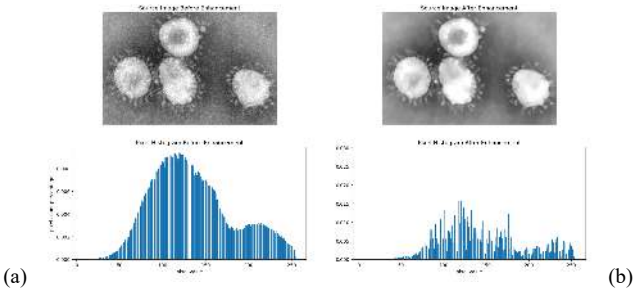
In this paper, original images and the preprocessed images are picked up for conducting our experiments. There are 6,000 electron microscopy images approximately which are equally split into four classes: SARS, MESR, HIV, and COVID-19. Each image has at least one region of interest (i.e., one bounding box contains one virus) which encapsulates visual objects with the same class, the predictive models referenced in this research support multiclass prediction for a given image. The data preprocessing methods in this research work are based on image augmentation as well as image quality enhancement.

A. Data Processing

The data preprocessing transformations are presented as the follows: Image denoising, image brightness adjustment, image contrast enhancement, image sharpening, image rotation, image random region removal, Jaccard index crop, image resizing, etc.

Provided the preprocessing for an input image, the first four steps are commenced for the betterment of image quality by using image enhancement, the electron microscopy images are likely noisy, followed by three image processing ways that aim to augment the source images by adding noises in order to improve the robustness of predictive model. The end of this preprocessing is to resize the images, that means, a predictive model requires multiple sizes of inputs.

The result of our image enhancement is shown in Fig. 1. The most significant operation is image denoising. Image contrast enhancement is apparent that the virus regions are seen with obvious increasement in luminance; in the histogram, the number of absolute white pixels surges dramatically. The contours of all virus regions are clear after sharpening, despite of being less visually observed.



(a) (b)

Fig. 1. An example of our data enhancement (a) before and (b) after

B. Prior Knowledge of Virus Morphology

In this paper, we put forward a novel R-CNN model, specifically designed for the recognition of four virus species, which includes well-known morphological attributes, e.g., spherical shape and virus regions surrounded by spike protein projections of coronavirus. In addition, it is noticeable that foreign objects might disturb the visual information of viruses, e.g., air bubbles.

In this project, we develop a mechanism that takes into account the visual data by adding an extra loss term to R-CNN. For example, if the predictor detects that an object is encapsulated in a bounding box without observing the main object surrounded by minor spherical objects, whilst predicting the object as a species of coronavirus, there is an extra loss to reflect this error. In this project, there are a number of recognizable morphological features of virus being identified and summarized to a comprehensive loss term.

This work has the mechanisms for object detection based on visual features of corona virus. The first one is primitively to identify the main body of the clipped image region (i.e., the largest region by taking use of white pixels) and the surrounding scattered white dots (ie.e., spike protein projections). The number of white dots is the metric that measures whether the enclosed object is a corona virus or not. The finding of isolated "white dots" is a typical issue for counting the number of connected components [47], which is resolved by using either depth first search (DFS) or breadth frist search (BFS).

There are a number of corona virus electron micrographs for inappropriately processed images, e.g., denoised images with a minor number of white dots, given the count c of $areaProj$, a loss term L_1 is

$$L_1(c) = \begin{cases} 2, & c = 0 \\ \frac{1}{c}, & 0 < c < 5 \\ 0, & c \geq 5 \end{cases}. \quad (1)$$

The second metric is the geometrical location of the white dots associated to the main body. All white-pixel areas are treated as polygons, whose centroids are calculated. Euclidean distances between the centroids to the main centroid are calculated. The centroid is the arithmetic mean of all points within this polygon [48].

Fig. 2. A region of interest from a coronavirus electron micrograph with a circular enclosure

Pertaining to the Euclidean distance, the standard deviation σ is computed as a heuristic to judge how the surrounding white dots are distributed. The greater σ is, the less likely white dots are surround main body. The second loss term is expressed as

$$L_2(\sigma) = k\sigma^3 \quad (2)$$

where we take use of a cubic power to amplify penalty so as to enlarge σ and k for controlling this loss term.

The third metric is a similarity. Given the centroid of a virus shape, a circle expands from the centroid point till the

contour of this virus is enclosed completely. The loss term is the ratio of two areas: $mainBody$ and $circleArea$, as $r = mainBody/circleArea$. Weight k and nonlinearity penalty item with a cubic power function are applied

$$L_3(r) = k\left(1 - \frac{1}{r}\right)^3. \quad (3)$$

However, the shape of a given virus is unlikely of a perfect circle in an electron microscopy image, there should be no penalty on a roughly round object. Thus, based on sampled virus images, a threshold r is set to relieve penalty for normal roughly round objects,

$$L_3(r) = \begin{cases} L_3(r), & r \geq 0.3 \\ 0, & r < 0.3 \end{cases}. \quad (4)$$

The fourth loss term is the ratio between width w and height h. Considered virus regions are roughly spherical, which usually turn up in electron microscopy images with a shape of a circle (contour), the bounding box should be a square rather than a rectangle. Hence, the loss term is expressed as

$$L_4(w,h) = k\left(\frac{\left|1-\frac{w}{h}\right| + \left|1-\frac{h}{w}\right|}{2}\right), \quad (5)$$

where k is a scalar that controls the importance of the penalty. Akin to the third loss term that there barely exists a perfect spherical virus particle, the penalty on normal approximately round objects should be avoided. Given this consideration,

$$L_4(w,h) = \begin{cases} L_4(w,h), & \left|1-\frac{w}{h}\right| \geq 0.2 \\ L_4(w,h), & \left|1-\frac{h}{w}\right| \geq 0.2 \\ 0, & otherwise. \end{cases} \quad (6)$$

We summarize all aforementioned loss terms with weight w, a loss term $L_{coronaV}$ (•) related to our prior knowledge is derived from

$$L_{coronaV}(y', y, t') = w_1 L_1 + w_2 L_2 + w_3 L_3 + w_4 L_4 \quad (7)$$

where y' and y are predicted object class and t' is the predicted bounding box that gives information in terms of the shape and anchor of a rectangular region in the image.

C. Attention Mechanism

A typical attention unit learns three weight matrices: Query weight W_Q, key weight W_K , and value weight W_V. For every input token x_i, there are $q_i = x_i W_Q$, $k_i = x_i W_K$, $q_i = v_i W_V$ for query and key value x_i. Attention weight a_{ij} from token x_i to token x_j is the dot product of q_i and k_j. We denote d_k as the number of dimension of q_i , and $a_{ij}/\sqrt{d_k}$ is the normalized attention weight. Thus, we define $Q = \{q_1, q_2, \dots, q_n\}$, $K = \{k_1, k_2, \dots, k_n\}$, $V = \{v_1, v_2, \dots, v_n\}$, the attention layer is represented as

$$attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (8)$$

where softmax(·) is the softmax function.

Given the example of an electron microscopy image of four corona virus regions, in order to train an attention layer, the label for each pixel and the surroundings of the pixel are shown in Figure 3. Regions of interest are labelled in white while backgrounds or the irrelevant are black. The desired output from the attention mechanism is the regions of interest. In practice, the attention mechanism cannot ensure to always find the optimal virus regions. Instead, in this experiment, soft mask is selected to filter out irrelevant regions in the source image [47,48].

(a) (b) (c)

Fig. 3. An electron microscopy image (a) with four corona virus regions, their masks (b) and the soft masked image (c) by utilizing attention mechanism

The proposed attention mechanism in this project is a simple self-attention network [49,50]. The network is regarded as a simple encoder-decoder network that only imports the source images, so that irrelevant regions are filtered out while the regions of interest are highlighted or preserved for visual feature extraction [49].

D. Evaluation Methods

In this paper, mAP, namely, mean average precision, is selected for evaluating multiple classifiers. Given a precision-recall curve, by incrementing a true-false threshold, mAP is found by using average precision. Weights are assigned to the samples of different labels so that imbalance samples with respect to all labels have defined proper weights (this operation is referred as "macro mode").

$$mAP = \frac{1}{Q}\sum_Q AveP(q) \quad (9)$$

where Q is the number of queries related to average precision score.

Recall is chosen to measure sensitivity of the classifications that a relevant record is retrieved from a set of queries. Similar to precision, we define a query set $Q = \{q_1, q_2, \dots, q_n\}$ and a relevant record set $R = \{r_1, r_2, \dots, r_m\}$, the computation is given as

$$Recall = \frac{count(R \cap Q)}{count(R)}. \quad (10)$$

F_1 score is given by eq. (10)

$$F_1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}, \quad (11)$$

where $Precision$ and $Recall$ are derived from the aforementioned equations. F_1 score is summarized as the harmonic mean of precision and recall [50], which is selected along with precision and recall to evaluate classification performances of the developed classifiers.

In order to compare different methods fairly, in this project, IOU is utilized to evaluate the proposed bounding boxes against the ground truth of bounding boxes. A proposed bounding box compares with all ground truth, the highest IOU percentage is selected as the IOU of the bounding box. The mean percentage of all bounding boxes produced by a regressor is regarded as the benchmark of bounding boxes. A bounding box only takes into account the overlapping areas; thus, any classification error is not reflected. Algorithm (1) shows this evaluation process.

Algorithm 1: Find bounding box for a regressor

Input: $\widehat{B_i}$ //a proposed bounding box of image I_i
B_i // ground truth of bounding box of image I_i
I // image set

Output: iou_mean // mean IOU of all proposed bounding boxes yielded by one regressor

$iou = \emptyset$
For I_i in I
  For $\widehat{B_{i,j}}$ in $\widehat{B_i}$ Then
    $iou_{B_{tmp}} = 0$
    For $B_{i,j}$ in B_i Then
      $iou_{B_{tmp}} = max\left(IOU(\widehat{B_{i,j}}, B_{i,j}), iou_{B_{tmp}}\right)$
    $iou = iou.append(iou_{B_{tmp}})$
$iou\_mean = average(iou)$

In this project, the losses of all predictive models are summed to evaluate the performance of the global model. In addition, the loss reflects the contribution of the prior morphological knowledge. In order to illustrate the training loss changes while gradients are gradually vanishing during the training, the losses are transformed in logarithm to indicate this trend.

IV. Experimental Results and Resultant Analyis

Our experiment employs R-CNNs to classify the type of viruses. Our predictive models and corresponding mAPs for the four virus species are shown in Table I. Despite mAPs are closer, the mAP of HIV virus detection is slightly higher than that of other four virus species. It should be noted that all predictive models illustrate mAPs up to 93%. Faster R-CNN performs better than its predecessors.

In this project, the pretrained results are displayed in mAP for each object class. All base networks show mAPs up to 91%. The results are resembled to that of classifiers, HIV virus detection is relatively higher than other base networks. The details are demonstrated in Table II.

The means of IOUs and MSE for bounding box areas are calculated for evaluating the performances of the predictive models. SSD is the most beneficial regarding the mean of IOUs, while Faster R-CNN has the most mediocre MSE_{bnd}. There is an inconspicuous discrepancy regarding both mean of IOUs and MSE of bounding boxes.

TABLE I. A SUMMARY OF CLASSIFICATION RESULTS (MAP) PER CLASSIFIER PER CLASS

	SARS	MERS	COVID-19	HIV
R-CNN	93.32 %	93.45 %	93.21 %	94.32 %
Fast R-CNN	93.41 %	93.31 %	93.37 %	94.71 %
Faster R-CNN	94.11 %	94.63 %	94.45 %	95.43 %
SSD	94.08 %	94.58 %	94.27 %	95.18 %
YOLO	93.92 %	94.93 %	94.32 %	95.01 %

TABLE II. A SUMMARY OF CLASSIFICATION RESULTS (MAP) FOR THE FIVE BASE NETWORKS

	SARS	MERS	COVID-19	HIV
VGG11	92.94 %	92.95 %	92.38 %	93.47 %
VGG16	93.12 %	93.51 %	93.14 %	93.94 %
GoogLeNet	93.51 %	93.33 %	93.72 %	93.96 %
ResNet152	93.62 %	93.15 %	93.78 %	93.69 %
AlexNet	92.22 %	92.38 %	92.25 %	92.71 %

TABLE III. THE MEANS OF IOUS FOR VARIOUS PREDICTIVE MODELS

	Means	*MSE*s
R-CNN	83.54 %	0.141
Fast R-CNN	84.91 %	0.134
Faster R-CNN	84.26 %	0.133
SSD	85.38 %	0.135
YOLO	83.83 %	0.138

TABLE IV. THE LOSSES (×100) AGAINST THE PRIOR KNOWLEDGE

	SARS	MERS	COVID-19	HIV
R-CNN	0.381	0.382	0.380	0.342
Fast R-CNN	0.421	0.423	0.421	0.431
Faster R-CNN	0.410	0.409	0.410	0.473
SSD	0.395	0.397	0.394	0.352
YOLO	0.432	0.431	0.433	0.409

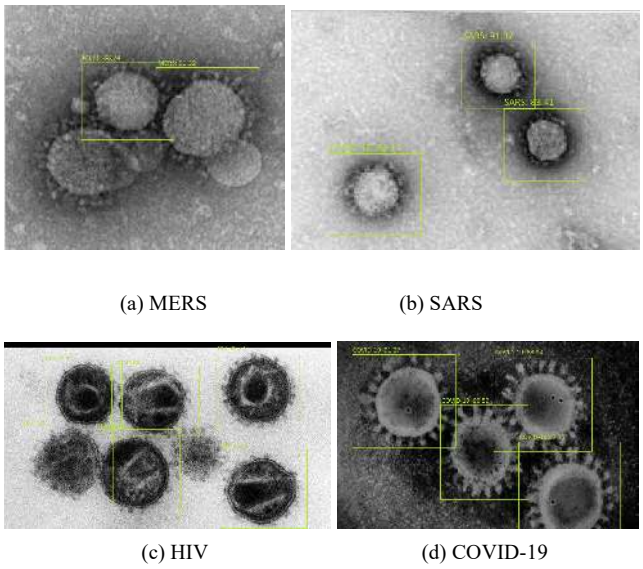
TABLE V. THE RESULTS BY UTILIZING VARIOUS UPDATING SCHEMES, MEASURED BY USING TOTAL LOSS

	Constant	Step-based decay	Time-based decay	Exponential decay
R-CNN	0.052	0.040	0.041	0.047
Fast R-CNN	0.035	0.029	0.027	0.028
Faster R-CNN	0.037	0.030	0.029	0.031
SSD	0.036	0.028	0.029	0.031
YOLO	0.038	0.029	0.027	0.032

In this project, the particular loss reflects the error $L_{coronaV}$ against the defined prior knowledge which is recorded per classifier per class. Given the minor weight on $L_{coronaV}$, the loss for $L_{coronaV}$ is multiplied by 100.0 for display. The total loss obtained through different learning rates are recorded in Table V, arranged by per method per classifier. The losses are verified by using validation dataset. R-CNN eyes that the greatest errors regardless of the choice of training rate schemes, in contrast to that of other predictive models with observable improvements in terms of loss values by using all kinds of learning rate schemes. Between different updating scheme of learning rates, the constant is the least favorable with greater loss values than the rest of schemes. There is not significant superiority of any particular schemes among the rest three (step-based decay, time-based decay, and exponential decay).

Figure 4 shows various results of virus detection in digital images. Figure 4(a) is a MERS image where three MERS virus regions are presented but only two are detected. The third object is low in illumination and the morphological features are vague. Figure 4(b) shows that the bounding boxes of ground truth should have been assigned with SARS labels. It is noticeable that the bounding boxes for the virus contain relatively large irrelevant areas that lead to wrong recognition by utilizing the detector while considering more significant areas as regions of interest. Figure 4(c) is shown for HIV virus detection. For HIV virus regions in this image, the confidence score is showcased and the proposed bounding boxes tightly surround the virus. Through reviewing the image samples, there is not suspiciously long rectangular bounding box, which is a strong indicator of the prior

knowledge being successful. However, this may play a role in the failure of detecting virus regions with irregular shapes.



(a) MERS (b) SARS

(c) HIV (d) COVID-19

Fig. 4. Various results of viruses detection from digital images

Spike proteins are crucial morphological features for virus electron micrography classification and undistinguished spikes for resemblances to background noises. For highlighting the spike protein regarding more reliable prediction performance, in this project, we brought in numerous image-processing methods for data augmentation.

SARS and COVID-19 are both appeared with spike proteins that provide the power of being highly contagious. The two types of viruses are from the identical species (corona viruses), thus share a degree of similarity in morphology. Therefore, the classification results are less satisfactory based on HIV as it has little similarities with other two. The strong similarities to corona visual features are found in MERS as well.

In this paper, we notice the similar prediction capability of the three classifiers from the R-CNN family. Amongst the results, R-CNN, Fast R-CNN, and Faster R-CNN have a similar performance. The proposed Faster R-CNN does not display accuracy with a considerable gap than the other two. This may be due to the inadequate source images for our model training. If this experiment could assemble more high-definition images, there exists a substantial possibility of having more accurate prediction, the three predictors would differentiate themselves by using the same evaluation metrics.

V. Conclusion

Owing to virus morphology, in this paper, we propose a novel loss term to reflect our predictions. The loss function mainly focuses on surrounding protein projections appeared in electron microscopy images as “white dots” weigh roughly spherical physical shape, which turns up in electron microscopy images as an approximately round polygon. The bounding box is taken into account which is constrained by an aspect ratio. However, the loss term is not heavily weighted. Consequently, a predictive model effectively manages the classification and regression of bounding boxes.

In this paper, we manipulate a total of five deep learning models: R-CNN, Fast R-CNN, Faster R-CNN, SSD, and YOLO. Amongst them, Faster R-CNN and SSD contributed to the most reliable results in terms of classification and bounding box regression. However, different models demonstrate various performance, there is not a particular network outperformed across all evaluation metrics. In this project, we evaluate the training processes of various networks by measuring loss convergence, all networks demonstrate the strong trend of convergence.

Our future work includes extensive considerations of virus placements so that deep learning models can recognize viruses amongst complex foreign objects [51]. We will propose a virus detection algorithm that distinguishes desired virus from image regions having noises. Thus, a large collection of electron microscopy images is expected [52].

References

[1] Williams, R. C. The shapes and sizes of purified viruses as determined by electron microscopy. In Springer Symposia on Quantitative Biology, vol. 18, pp. 185-195, 1953.

[2] Bachrach, H. L., Breese Jr, S. S. “Purification and electron microscopy of foot-and-mouth disease virus, ” In Proceedings of the Society for Experimental Biology and Medicine, 97(3), 659-665, 1958.

[3] Matuszewski, B. J., Shark, L. K. Hierarchical iterative Bayesian approach to automatic recognition of biological viruses in electron microscope images. In International Conference on Image Processing, Vol. 2, pp. 347-350, 2001.

[4] Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., Garry, R. F. “The proximal origin of SARS-CoV-2, ” Nature Medicine, 26(4), 450-452, 2020.

[5] Lu, X., Zhang, L., Du, H., Zhang, J., Li, Y. Y., Qu, J., Wu, C. “SARS-CoV-2 infection in children,” New England Journal of Medicine, 382(17), 1663-1665, 2020.

[6] Hatcher, W. G., Yu, W. “A survey of deep learning: Platforms, applications and emerging research trends, ” IEEE Access, 6, 24411-24432, 2018.

[7] Krizhevsky, A., Sutskever, I., Hinton, G. E. ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems, pp. 1097-1105, 2012.

[8] Lim, W., Jang, D., Lee, T. Speech emotion recognition using convolutional and recurrent neural networks. In APSIPA, pp. 1-4, 2016.

[9] Komar, M., Yakobchuk, P., Golovko, V., Dorosh, V., Sachenko, A. Deep neural network for image recognition based on the Caffe framework. In IEEE International Conference on Data Stream Mining & Processing (DSMP), pp. 102-106, 2018.

[10] Hatami, N., Gavet, Y., Debayle, J. Classification of time-series images using deep convolutional neural networks. In ICMV 2017, pp. 106960Y.

[11] Termritthikun, C., Kanprachar, S. Accuracy improvement of Thai food image recognition using deep convolutional neural networks. In International Electrical Engineering Congress, pp. 1-4, 2017.

[12] Chen, S., Liu, Y., Gao, X., Han, Z. Mobilefacenets: Efficient CNNs for accurate real-time face verification on mobile devices. In Chinese Conference on Biometric Recognition, pp. 428-438, 2018.

[13] Cui, Q., McIntosh, S., Sun, H. “Identifying materials of photographic images and photorealistic computer generated graphics based on deep CNNs, ” Comput. Mater. Continua, 55(2), 229-241, 2018.

[14] Alvarez-Borrego, J., Chavez-Sanchez, M. C. “Detection of IHHN virus in shrimp tissue by digital color correlation, ” Aquaculture, 194(1-2), 1-9, 2001.

[15] Wang, L., Eng, E. T., Law, K., Gordon, R. E., Rice, W. J., Chen, B. K. “Visualization of HIV T cell virological synapses and virus-containing compartments by three-dimensional correlative light and electron microscopy, ” Journal of Virology, 91(2), 2017.

[16] Lin, Y., Yan, X., Cao, W., Wang, C., Feng, J., Duan, J., Xie, S. “Short communication probing the structure of the SARS coronavirus using scanning electron microscopy, ” Antiviral Therapy, 9, 287-289, 2004.

[17] Yuan, Y., Cao, D., Zhang, Y., Ma, J., Qi, J., Wang, Q., Zhang, X. “Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains, ” Nature Communications, 8(1), 1-9, 2017.

[18] Prasad, S., Potdar, V., Cherian, S., Abraham, P., Basu, A., Team, I. N. N. “Transmission electron microscopy imaging of SARS-CoV-2, ” The Journal of Medical Research, 151(2-3), 241, 2020.

[19] Gilbert, P., McKeague, I., Eisen, G., Mullins, C., Guéye-NDiaye, A., Mboup, S., Kanki, P. “A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening,” Statistics in Medicine, 22(4), 573–593, 2003.

[20] McGovern, S. L., Caselli, E., Grigorieff, N., Shoichet, B. K. “A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening,” Journal of Medicinal Chemistry, 45(8), 1712-1722, 2002.

[21] Wong, C., Li, X., K., Lau, S., C., Woo, P. “Global epidemiology of bat coronaviruses,” Viruses, 11(2), 174, 2019.

[22] Goldsmith, C. S., Tatti, K. M., Ksiazek, T. G., Rollin, P. E., Comer, J. A., Lee, W. W., Zaki, S. R. “Ultrastructural characterization of SARS coronavirus, ” Emerging Infectious Diseases, 10(2), 320, 2004.

[23] Groot, R. J., Baker, S. C., Baric, R. S., Brown, C. S., Drosten, C., Enjuanes, L., S, E. J. “Middle East respiratory syndrome coronavirus (MERS-CoV): Announcement of the coronavirus study group, ” Journal of Virology, 87(14), 7790–7792, 2013.

[24] Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Hilgenfeld, R. “Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors,” Science, 368(6489), 409-412, 2020.

[25] Girshick, R., Donahue, J., Darrell, T., Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In International Conference on Computer Vision and Pattern Recognition, pp. 1-21, 2017.

[26] Wang, X., Yang, M., Zhu, S., Lin, Y. RegionLets for generic object detection, International Conference on Computer Vision, pp. 17-24, 2013.

[27] Uijlings, J., Sande, K. v., Gevers, T., Smeulders, A. “Selective search for object recognition, International Journal of Computer Vision, 104(2), 154-171, 2013.

[28] Comaniciu, D., Meer, P. “Mean shift: A robust approach toward feature space analysis, ” IEEE Transactions on Pattern Analysis and Machine Intelligence, 24, 603-619, 2002.

[29] Chen, C., Liu, M. Y., Tuzel, O., Xiao, J. R-CNN for small object detection. In Asian Conference on Computer Vision, pp. 214-230, 2016.

[30] Girshick, R. “Fast R-CNN,” International Journal of Computer Vision, pp. 1440-1448, 2015.

[31] Gua, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B. S., Chen, T. “Recent advances in convolutional neural networks,” Pattern Recognition, 77, 354-377, 2018.

[32] He, K., Zhang, X., Ren, S., Sun, J. “Spatial pyramid pooling in deep convolutional networks for visual recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(9), 1904-1916, 2015.

[33] Krizhevsky, A., Sutskever, I., Hinton, G. ImageNet classification with deep convolutional neural networks. International Conference on Neural Information Processing Systems, pp. 1097-1105, 2012.

[34] Ren, S., He, K., Girshick, R., Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. International Conference on Neural Information Processing Systems, pp. 91-99, 2015.

[35] Long, J., Shelhamer, E., Darrell, T. Fully convolutional networks for semantic segmentation. IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431-3440, 2015.

[36] Chorowski, J. K., Bahdanau, D., Serdyuk, D., Cho, K., Bengio, Y. Attention-based models for speech recognition. Neural Information Processing Systems, pp. 577-585, 2015.

[37] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. You Only Look Once: Unified, real-time object detection. ICCV, pp. 779-788, 2015.

[38] Lin, T. Y., Roy, A., Maji, S. Bilinear CNNs for fine-grained visual recognition, 2015 ArXiv:1504.07889.

[39] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A. C. SSD: Single shot multibox detector. ICCV, pp. 21-37, 2015.

[40] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Li, F.-F. “ImageNet large scale visual recognition challenge,” International Journal of Computer Vision, 115(3), 211–252, 2015.

[41] Chen, S., Hong J., Zhang, T, Li, J., Guan, Y. Object detection using deep learning: Single shot detector with a refined feature-fusion structure. IEEE International Conference on Real-time Computing and Robotics (RCAR), 2019.

[42] Tang, X., Du, D. K., He, Z., Liu, J. PyramidBox: A context-assisted single ShotFace detector. ECCV, pp. 1-17, 2018

[43] Kaur, M. A., Sharma, M. P., Verma, M. A. A appraisal paper on breadth-first search, depth-first search and Red black tree. International Journal of Scientific and Research Publications, 4(3), 2-4, 2014.

[44] Barton, A. College Calculus with Analytic Geometry. Addison-Wesley, 1972.

[45] Ramachandran, P., Parmar, N., Vaswani, A., Bello, I., Levskaya, A., Shlens, J. Stand-alone self-attention in vision models. International Conference on Neural Information Processing Systems, pp. 68-80, 2019.

[46] Hand, D., Christen, P. “A note on using the F-measure for evaluating record linkage algorithms, ” Statistics and Computing, 28(3), 539-547, 2018.

[47] Pan, P., Yan, W. “Object detection based on saturation of visual perception,” Multim Tools Appl. 79(27-28): 19925-19944, 2020.

[48] Pan, C., Li, X., Yan, W. A learning-based positive feedback approach in salient object detection. IEEE IVCNZ, pp. 1-6, 2018.

[49] Ji, H., Liu, Z., Yan, W., Klette, K. Early diagnosis of Alzheimer's disease based on selective kernel network with spatial attention. ACPR pp. 503-515, 2019.

[50] Ji, H., Yan, W., Klette, R. Early diagnosis of Alzheimer’s disease using deep learning. ICCCV, pp. 87–91, 2019.

[51] Al-Sarayreh, M., Reis, M., Yan, W., Klette, R. A sequential CNN approach for foreign object detection in hyperspectral images. CAIP pp. 271-283, 2019.

[52] Zhang, L. Virus Identification From Digital Images Using Deep Learning, Masters Thesis, Auckland University of Technology, Auckland, New Zealand, 2020.