INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

www.ijraset.com

Call: ✆ 08813907089 | **E-mail ID:** ijraset@gmail.com

The Application of Focal Loss in various Domains: A Survey

Aishwarya Seth[1], Dr. B. Satish Babu[2]

[1, 2]Computer Science & Engineering Department, R.V. College of Engineering

Abstract: Focal loss is a novel metric introduced in the recent past to resolve class imbalance problem. This loss was found to have a wider impact in various factions of object identification research, and has been centred around improved object detection techniques. The applications of object detection techniques are wide and varied, and the introduction of focal loss as a deterministic metric during training has had a revolutionary impact on the research community. This paper surveys the introduction and resulting impact of this new metric to various applications.
Keywords: Focal loss, object detection, metric, survey, applications

I. INTRODUCTION

Object detection techniques have shown pivotal outcomes for algorithmic research in the field of artificial intelligence, particularly in the case of neural networks. Currently, majority of the newer and more advanced techniques use several cascading layers, involving multiple stages, with some performing the task of feature extraction while others focus on the proposals of regions of interest. These approaches have been viable and have resulted in good and stabilized, accuracy values. However, the multi-stage approach is usually computationally very expensive, and as a result, requires advanced hardware for its effective use and deployment.

In 2017, Facebook's AI Research team created a new loss metric, known as focal loss [1]. This metric was designed to be used during training of neural networks. The approach was aimed at improving the performance of one-stage object detection. Further, most prominent object detection algorithms today use multi-stage methods, which are complex neural networks and are also computationally expensive. To reduce dependency, the focal loss metric aims at shifting the focus to the hard, misclassified examples, and thereby improves the performance of one-stage detectors, without a subsequent significant change in their complexity. Thus, this represents a method to achieve a good accuracy without relying on multi-staged networks.

II. IMPORTANCE OF THE DEVELOPMENT

Majority of datasets for a variety of applications have some form of class imbalance. This is because it is nearly impossible to account for all the special conditions or corner cases that may exist in real-world data. Models that are designed by developers need to be robust, and must be able to gracefully tolerate all possible edge cases.

However, since data for edge cases is difficult to gather, the model may not behave correctly for certain classes. This leads to a severe imbalance in the data that is collected to tackle any particular problem statement. For example, if the user aims to classify words, there are short words like 'and' and 'the' that may be very common while other words like 'uncanny' are used sparingly. This class imbalance problem occurs frequently in the real-world, and makes it difficult to design classifiers that are able to respond to a large variety of situations. Additionally, this skew in the data frequently leads to high values of false positives or false negatives. To improve the effects of recursive class imbalance, two stage detectors were designed. These focused on the feature maps and were able to provide some tolerance. However, these methods are computationally demanding and are highly complex systems. To tackle the issue of the class imbalance problem, the focal loss was introduced.

This metric allows a network to understand the complexity of the current sample and provides a weight to each training sample based on its rarity. Thus, simple single stage networks are able to tackle the class imbalance problem by the incorporation of focal loss.

Since its first release in 2017, this metric has impacted a wide variety of applications in various fields. The impact of this metric across the fields of its applications has been studied in detail in this research work.

Section 3 covers the theory of this approach, section 4 deals with various domains of usage of the focal loss technique. Section 5 covers the analysis of results by using focal loss, Section 6 covers some possible future work, before section 7 covers the conclusion of this metric.

III. ABOUT FOCAL LOSS

Various forms of loss are used to measure the performance of a classification model, and to depict how the loss varies over time. One prominent example is the cross-entropy loss, which focuses on a reduction between the distances in two probability distributions, namely the actual and the predicted distributions. The primary flaw is that the cross entropy loss is dependent on the distribution of the input data. If one class has many more examples than the other, there will be an inherent imbalance in the probability distribution curves, which in turn has a detrimental effect on the ability of the designed classifier to tolerate rare cases.

To improve this class imbalance effect, the focal loss was introduced. Modifying the existing standard cross-entropy equation, focal loss introduced a factor represented as:

$$(1-p_t)^\gamma \tag{1}$$

By setting a positive gamma value, the focus of the neural network that is being trained shifts from the already well-classified examples to the misclassified, well-defined examples.

To begin with, the parameter p_t is a notation for the probability of the class labelled with numeric value 1. If the variable y represents the ground-truth value and is either 0 or 1, the probability of $p \in [0,1]$ is the probability that is estimated by the model for the ground truth or actual class of $y = 1$. In order to attain some notational ease, the variable p_t was defined as:

$$p_t = \begin{cases} p & \text{if } y=1 \\ 1-p & \text{otherwise} \end{cases} \tag{2}$$

The experiments that had been performed on various samples using the pre-existing cross entropy loss indicated that the results were easily skewed by having a large class imbalance. Some forms of balancing could be applied by introducing the balancing factor known as α, which did manage to cause a differentiation between the positive and negative samples. However, this approach was not able to balance the difference between the easy and hard samples.

The focal loss is calculated as follows:

$$FL(p_t) = -(1-p_t)^\gamma \log(p_t) \tag{3}$$

The modulating factor $(1-p_t)^\gamma$ is thus added to the original cross entropy formula. When an instance passed to the model is misclassified and p_t is small, the modulating factor is nearly unity and thus does not affect the loss. As p_t tends to 1, the modulating factor tends to 0, and again, the loss is not influenced. Thus, all well-classified examples have lower weightage. This down weighting of the various examples can be controlled smoothly by the modulating factor, γ.

IV. APPLICATIONS OF FOCAL LOSS FOR OBJECT DETECTION

A. Object Detection on Roads

1) *Vehicle Detection*: The introduction of focal loss in vehicles began with its application to surveillance of vehicles. [2] It was observed that the RetinaNet based focal loss training that was performed on the UA-DETRAC dataset [3] was able to find the vehicles appropriately even in various and varied light conditions and across frames. It was thus an encouraging result for extension of the RetinaNet for vehicle detection in intelligent transport systems. This focal loss based approach rivalled the accuracy of the two-stage Faster RCNN approach, providing a much faster solution with a comparable accuracy metric. The inferencing for this particular set of applications is usually done on the multi-object tracking dataset created by the University of Albany, known as the DEtection and TRACking (DETRAC) dataset. [4].

2) *Traffic Light Detection:* Another important aspect of vehicles which is strongly influenced by artificial intelligence is the advanced driver assistance system (ADAS). To detect smaller traffic lights effectively, this metric of loss was incorporated in deep neural networks. [5] The method used in this paper implemented a focal regression loss. The results were a higher mAP than state-of-the-art detectors by 7.19 to 42.03% on the Bosch-TL dataset [6] and by 19.86 to 49.16% on the LISA-TL dataset [7] The focal regression approach was also extended to general object detection techniques to test its capability of improving the class imbalance effects. [8] The introduction of a γ was done. This value is dependent on the progress made by the training and thus it achieves a better performance. To reduce the overhead computation of the focal loss hyperparameters, automated focal loss has been introduced by this research work. This particular method has been applied primarily to the task of detecting vehicles in three-dimensional spaces.

B. Object Detection in Images and Videos

1) *Image Captioning*: The introduction of a focal cross-entropy loss was found to be beneficial while detecting the various attributes of a pedestrian and returning them as part of the caption for a particular input image. [9] By introducing this novel fusion attention mechanism which is based on focal loss, generative adversarial networks can be improved. A similar metric was introduced for small object detection within a video frame, which is known as the ω focal loss. [10] This allows for small and dense detections in images, enabling captioning of the information. Automatic image tagging [11] can also be achieved using transfer learning in deep networks with the focal loss metric applied. The introduction of focal loss in vehicles began with its application to surveillance of vehicles. [2] It was observed that the RetinaNet based focal loss training that was performed on the UA-DETRAC dataset [3] was able to find the vehicles appropriately even in various and varied light conditions and across frames. It was thus an encouraging result for extension of the RetinaNet for vehicle detection in intelligent transport systems. This focal loss based approach rivalled the accuracy of the two-stage Faster RCNN approach, providing a much faster solution with a comparable accuracy metric. The inferencing for this particular set of applications is usually done on the multi-object tracking dataset created by the University of Albany, known as the DEtection and TRACking (DETRAC) dataset. [4].

2) *Image Component Recognition:* A combination of the connectionist temporal classification along with focal loss has been used to create a more advanced optical character recognition system, particularly for Chinese characters. [12] Thus, being passed an input image, this algorithm returns the composing characters. In a similar fashion, the creation of multi-class labels in case of a single-label input was found to create a more generalized output and with accuracy metrics that rival those of single-label predictors. [13] In a niche avenue like image captioning, which frequently has sparse samples for training, the usage of focal loss helps improve the performance of neural networks designed for this task.

C. Object Detection in Medical Images and Samples

1) *Segmentation and Area Captioning*: In the segmentation and captioning of medical images, due to the data imbalance, this metric for loss estimation is utilized. [14] Automatic detection of skin lesions is also strongly supported by a combination of loss metrics, specifically a combination of the focal loss and the jaccard distance measures. [15] The approach was also extended to the detection of ischemic stroke lesion detections [16] and the primary role of the focal loss metric was to prevent the increase in the false positives rate. A hybrid approach with hybrid focal loss measurements has been employed for this particular detection. Apart from these specific classification techniques, object instance segmentation and region of interest detection is done via the focal loss metric by implementing a variation of the Mask RCNN algorithm. This method has been implemented to perform the segmentation of lung nodules. [17]

2) *Classification of Medical Samples:* An improvement on existing methods of testing whether a breast cancer sample is benign or malignant was achieved when a multi-scale ResNet structure was utilized along with the focal loss metric for training of samples. [18]. Similar classifications of lung nodules have been performed by using convolutional neural networks with focal loss. [19] After the preliminary separation into nodule and non-nodule, the categorization of the lung nodules incorporates a specifically designed neural network with focal loss as the convergence metric. In order to overcome the few numbers of abnormal red blood cell samples and to automate the classification of their morphologies, focal loss based convolutional neural networks were used. [20] This is a high impact application, since it requires pathologists to be experts in the domain and misdiagnosis leads to improper treatment. This pressure is alleviated by the use of neural networks for the classification task. Thus, the focal loss metric has had a significant impact on the development and improvement of applications in the biomedical domain. This is expected since most samples in this particular domain are generally primarily constricted by the skew present in the input data, that is, an imbalance in the available samples for each class.

D. Object Detection for People and Face Detection

1) *Detection of People*: Focal loss metrics have been used in person re-identification from video surveillance streams. [21] This approach utilized spatial attention and temporal regularization while performing the re-identification of a particular subject in question. Using focal and L1 loss for classification and bounding box regression respectively, the person detection subnet is improved during the construction of the MultiPoseNet, a keypoint-based, multiple person segmentation algorithm. [22] This form of loss has been extended to pedestrian detection mechanisms by using asymptotic localization filtering to incrementally improve the bounding boxes of each of the proposed anchors. [23]

2) *Face Detection and Recognition:* The approach has been extended to the face recognition system, including a DenseNet based face detection framework [24]. By usage of this metric, attention networks have been created, which are also able to detect partially visible faces. [25] This is particularly useful to break down images of crowded environments. The alignment of the face recognition was accomplished by focal loss based multi task convolutional neural networks. [26] The success of this endeavor has been so promising that the focal loss metric is also used in creating real-time face detectors that are ported and operational on embedded systems. [27]

E. Other Applications

Some other interesting applications of the focal loss method have been in an adaboost and convolutional neural network based drogue detection during autonomous refueling. [28] The resulting algorithm showed improved performance in terms of the accuracy as well as timing metrics. Neural graph filtering can be applied to fashion recommendation systems, and the focal loss based metric resulted in an improvement by nearly 10% in comparison with the earlier methods. The design also had a high customer preference rating. [29] With an increasing importance on biometrics, the introduction of focal loss to establish a one-stage effort for fingerprint singular point detection is bound to have an impact on security efficiency in various applications everywhere. [30] The false alarm rate and detection speed are also influenced positively. The focal loss technique has also been applied to agriculture-based applications, where it has been used to detect diseases within the Casava plant. [31] The paper managed to achieve an accuracy of 93% in their detection of the various diseases, and thus had an impactful outcome in solving a large-scale problem.

During autonomous driving or in data collection about roads in an area, detection of whether they are paved or unpaved is critical to providing suggestions to the driver. This particular form of detection can be improved significantly by the introduction of the focal loss metric during the training of the associated algorithm. [32] By grouping and detecting corners, object detection can be performed. [33] In order to improve this process, a variant of focal loss was proposed. This variation of the metric is designed to reduce the penalty that is given to negative locations that are found within some specific distance from a positive location. To perform semantic segmentation, the designed network should be able to differentiate among the unique composite features. While learning these features, one particular method of providing guidance is to predict the boundary of the semantic region. [34] This method is guided by the changes that take place in the focal loss, which acts as a supervisor for the output of this section of the network. From a raw point cloud, an entire three-dimensional object can be detected from a point-based region convolutional neural network. [35] Here, the number of points that are found in the foreground place usually exceeds the number of points that form the background. To improve the imbalance that occurs because of this skew in point density, the focal loss metric has been utilized. Thus, due to its inherent class imbalance tolerance, the focal loss has a number of varied applications.

V. ANALYSIS AND COMPARISON OF RESULTS ACROSS APPLICATIONS

A. Evaluation Metrics

1) *Mean Average Precision:* The primary parameter that is used for comparison of the improvement provided by the focal loss technique is known as the mean average precision. (mAP) The value of precision for a particular class is the ratio of true positives to total samples. The average precision is then calculated for all the classes and the average of these values is taken as the final mAP value. The equation used for this calculation is:

$$mAP = \frac{\sum_{q=1}^{Q} AveP(q)}{Q}$$

(4)

Where, Q is the number of queries in a given set and AveP(q) is the average precision value calculated for a particular user entry q. The mAP metric was developed particularly for usage in object detection techniques. This metric can also be evaluated as the area under the precision recall curve. A basis for estimating mAP across various scales is dependent on the intersection over union (IoU) metric. The results in the form of boundary boxes are selected according to different IoU thresholds in order to calculate the mAP. The IoU metric is defined as the ratio between the area of overlap and the area of union. It is a measure of how well the boundary box matches that of the ground truth. The mAP metric is usually represented in the form of a percentage, which is followed in the subsequence sections and discussions. In the segmentation and captioning of medical images, due to the data imbalance, this metric for loss estimation is utilized [14]. Automatic detection of skin lesions is also strongly supported by a combination of loss metrics, specifically a combination of the focal loss and the jaccard distance measures [15]. The approach was also extended to the detection of ischemic stroke lesion detections [16] and the primary role of the focal loss metric was to prevent the increase in the false positives rate. A hybrid approach with hybrid focal loss measurements has been employed for this particular detection.

Apart from these specific classification techniques, object instance segmentation and region of interest detection is done via the focal loss metric by implementing a variation of the Mask RCNN algorithm. This method has been implemented to perform the segmentation of lung nodules [17].

2) *Accuracy Metric:* The most common metric for basic classification is the accuracy metric. This is calculated by estimating the number of correct samples, divided by the total number of samples that are collected. In some cases, soft accuracy is utilized. This method allows some tolerance parameters to be introduced to the system, such as allowing an error of a step of 1. Here, a step is the number of edits needed for the prediction to match the ground truth [12].

3) *Miss Rate over False Positive Per Image on log scale*: This particular scale was proposed by Caltech and is a standard form of evaluation, [36] particularly designed for pedestrian detection. This method is referred to as MR in subsequent sections. Miss rate is defined as the number of false negatives out of the total of true positives and false negatives. This value is plotted against the number of false positives per image (FPPI) by varying the threshold of confidence of detection. It is used in pedestrian detection since there is usually found to be an upper limit to the acceptable FPPI which is independent of the person density in the input image. To find the log-average Miss Rate over FPPI, both the values are plotted on a log scale, with values ranging from 10^{-2} to 10^{0} at nine equally spaced out FPPI rates. The value is then averaged across all the samples collected to get the final metric. This metric is an indication of the objects that are not detected.

B. Object Detection on Roads

The comparison of the various object detection techniques done on the road are done on the basis of their mAP metric as captured in Table 1.

TABLE I
mAP values for various object detection on roads

Application	Model Used	Dataset	mAP
Vehicle surveillance	RetinaNet	DETRAC	73.79
Vehicle detection for intelligent transport	RetinaNet	DETRAC	61.2
Traffic light detection for ADAS	Deconvolutional Deep Neural Network	BOSCH-TL	42.03
Traffic light detection for ADAS	Deconvolutional Deep Neural Network	LISA-TL	49.16
Automated Focal Loss for General Object Detection	RetinaNet with IoU metric set to 0.5	COCO [37] benchmark dataset	51.18

These values indicate an improvement over the existing state-of-the-art single stage methods, and show that this approach can be used to work with imbalanced classes.

C. Object Detection in Images and Videos

The proposed IA^2 net [9] shows an improvement in the mAP metric when focal loss is introduced. When inferencing is done on the Richly Annotated Pedestrian dataset, [38] the mAP has a value of 77.44. For small and dense commodity detection, the designed CommodityNet shows an mAP metric of 80.7 when modelled with the ResNet 50 approach.

The inferencing of the focal loss based optical character recognition showed an accuracy of 76.4% on the general Chinese character dataset. [39] The soft accuracy which allowed for a shift by one character, had a value of 94.6% in this particular experiment.

D. Object Detection in Medical Images and Samples

The results of the classification methods that are applied to the medical samples are captured along with the designed architecture, as shown in table 2.

TABLE III

Accuracy Values for Object Detection in Medical Images

Classification Type	Model Used	Dataset	Accuracy
Lesion Classification	Semi-supervised learning with adversarial training in a convolutional neural network based on Xception modules	IDRID [40]	91.34
Breast Cancer Classification	ResNet-50	BreaKHis [41]	94.92 at 40X magnification
Lung Node Classification	Proposed convolutional network	LIDC/IDRI [42]	97.20
Canine Red Blood Cell Morphology Classification	DenseNet-121	Custom dataset	95.60

E. Object Detection for People and Face Detection

For detection of people using keypoints or other mechanisms, various metrics are used. By using accuracy metrics on the PRID 2011 [43] dataset for person reidentification, the network using spatial attention and temporal pooling methods achieved an overall accuracy of 92.1%. On performing the person detection on the COCO dataset and applying an IoU threshold of 0.5, the mAP result was 81.5%.

The MR metric applied for single stage pedestrian detection was at a level of 12 for the Reasonable subset of the CityPersons dataset. [44] It attained an encouraging value of 51.9 on the Heavy subset, 11.4 on the Partial subset and 8.4 on the Bare subset, thus showing a robust response to various person density inputs that are available in the dataset.

While performing face detection, the input categories are segregated into easy, medium, and hard subsets. The results of the face detection applications of focal loss are captured in Table 3.

TABLE IIIII

Average Precision (AP) Values for People and Face Detection

Application	Network	Dataset	AP (Easy)	AP (Medium)	AP (Hard)
Face Detection	Dual Shot Face Detector with focal loss	Wider Face [45]	96.9	96.2	92.5
Face Detection of Occluded Faces	Uses a face attention network	Wider Face	94.6	93.6	88.5

VI. FUTURE PROSPECTS

The application of deep learning and other neural network-based systems to nearly all fields of life is an essential part of the current developments in the technological world. While working with real-world data, a skew in the input data sample is inevitable. Focal loss represents a viable and scalable solution for dealing with complex problem statements.

Thus, the applications of single shot detectors can be extended to all forms of object detection. Currently, the research focuses on detection of people, objects on the road, or medical information. With an increase in the robust capabilities of these single shot detection systems, the design of viable systems will be sped up and enhanced considerably. The extension of object detection systems in three-dimensional, as well as the detection of minute objects that denote various attributes in images or in video frames can also be explored. This focal loss technique will be crucial in further development of artificial intelligence systems for medical systems, as these usually face data constraints.

VII. CONCLUSIONS

The focal loss is a new and powerful metric that can allow for an alleviation in the common problem of class imbalance. This metric has been implemented successfully across a variety of fields, and has had an important impact in the improvement of object detection techniques that focus on using only single stage classifiers. The impact of this approach in the deep learning framework is expected to grow exponentially in the near future.

REFERENCES

[1]   Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He and Piotr Dollár, "Focal Loss for Dense Object Detection", arXiv:1708.02002 [cs.CV], Aug. 2017

[2]   X. Wang, P. Cheng, X. Liu and B. Uzochukwu, "Focal loss dense detector for vehicle surveillance," 2018 International Conference on Intelligent Systems and Computer Vision (ISCV), Fez, 2018, pp. 1-5.

[3]   M. Bugeja, A. Dingli, M. Attard, and D. Seychell, "Comparison of Vehicle Detection Techniques applied to IP Camera Video Feeds for use in Intelligent Transport Systems," Transportation Research Procedia, vol. 45, pp. 971–978, 2020.

[4]   L. Wen, D. Du, Z. Cai, Z. Lei, M. Chang, H. Qi, J. Lim, M.-H. Yang, and S. Lyu, "Detrac: A new benchmark and protocol for multi-object tracking," arXiv preprint arXiv:1511.04136, 2015.

[5]   E. Lee and D. Kim, "Accurate traffic light detection using deep neural network with focal regression loss," Image and Vision Computing, vol. 87, pp. 24–36, Jul. 2019, doi: 10.1016/j.imavis.2019.04.003.

[6]   K. Behrendt, L. Novak, R. Botros, A deep learning approach to traffic lights: detection, tracking, and classification, IEEE International Conference on Robotics and Automation, 2017. pp. 1370–1377.

[7]   M.P. Philipsen, M.B. Jensen, A. Møgelmose, T.B. Moeslund, M.M. Trivedi, Traffic light detection: a learning algorithm and evaluations on challenging dataset, IEEE International Conference on Intelligent Transportation Systems, 2015. pp. 2341–2345.

[8]   Michael Weber, Michael Fürst and J. Marius Zöllner, "Automated Focal Loss for Image based Object Detection", arXiv:1904.09048 [cs.CV], Apr. 2019.

[9]   Z. Ji, E. He, H. Wang, and A. Yang, "Image-attribute reciprocally guided attention network for pedestrian attribute recognition," Pattern Recognition Letters, vol. 120, pp. 89–95, Apr. 2019, doi: 10.1016/j.patrec.2019.01.010.

[10] Z. Ji, Q. Kong, H. Wang, and Y. Pang, "Small and Dense Commodity Object Detection with Multi-Scale Receptive Field Attention," in Proceedings of the 27th ACM International Conference on Multimedia, 2019, doi: 10.1145/3343031.3351064.

[11] T.-H. Tran, X.-H. Tran, V.-T. Nguyen, and K. Nguyen-An, "Building an Automatic Image Tagger with DenseNet and Transfer Learning," in 2019 International Conference on Advanced Computing and Applications (ACOMP), 2019, doi: 10.1109/acomp.2019.00013.

[12] X. Feng, H. Yao, and S. Zhang, "Focal CTC Loss for Chinese Optical Character Recognition on Unbalanced Datasets," Complexity, vol. 2019, pp. 1–11, Jan. 2019.

[13] Qi Dong, Xiatian Zhu and Shaogang Gong, "Single-Label Multi-Class Image Classification by Deep Logistic Regression", arXiv:1811.08400 [cs.CV], May 2019

[14] Yi Zhou, Xiaodong He, Lei Huang, Li Liu, Fan Zhu, Shanshan Cui, Ling Shao; "Collaborative Learning of Semi-Supervised Segmentation and Classification for Medical Images", The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 2079-2088

[15] L. Song, J. P. Lin, Z. J. Wang, and H. Wang, "An End-to-end Multi-task Deep Learning Framework for Skin Lesion Analysis," IEEE Journal of Biomedical and Health Informatics, pp. 1–1, 2020, doi: 10.1109/jbhi.2020.2973614.

[16] S. Qamar, H. Jin, R. Zheng, and M. Faizan, "Hybrid loss guided densely connected convolutional neural network for Ischemic Stroke Lesion segmentation," in 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), 2019, doi: 10.1109/i2ct45611.2019.9033802.

[17] H. Yan, H. Lu, M. Ye, K. Yan, Y. Xu, and Q. Jin, "Improved Mask R-CNN for Lung Nodule Segmentation," in 2019 10th International Conference on Information Technology in Medicine and Education (ITME), 2019, doi: 10.1109/itme.2019.00041.

[18] W. Shu, S. Wang, Q. Chen, Y. Hu, Z. Cai, and R. Lin, "Pathological Image Classification of Breast Cancer Based on Residual Network and Focal Loss," in Proceedings of the 2019 3rd International Conference on Computer Science and Artificial Intelligence, 2019.

[19] G. S. Tran, T. P. Nghiem, V. T. Nguyen, C. M. Luong, and J.-C. Burie, "Improving Accuracy of Lung Nodule Classification Using Deep Learning with Focal Loss," Journal of Healthcare Engineering, vol. 2019, pp. 1–9, Feb. 2019.

[20] K. Pasupa, S. Vatathanavaro, and S. Tungjitnob, "Convolutional neural networks based focal loss for class imbalance problem: a case study of canine red blood cells morphology classification," Journal of Ambient Intelligence and Humanized Computing, Feb. 2020

[21] X. Wang and X. Zhao, "Temporal Regularized Spatial Attention for Video-Based Person Re-Identification," in 2019 IEEE International Conference on Image Processing (ICIP), 2019.

[22] Muhammed Kocabas, Salih Karagoz, Emre Akbas; "MultiPoseNet: Fast Multi-Person Pose Estimation using Pose Residual Network" The European Conference on Computer Vision (ECCV), 2018, pp. 417-433

[23] Wei Liu, Shengcai Liao, Weidong Hu, Xuezhi Liang and Xiao Chen; "Learning Efficient Single-stage Pedestrian Detectors by Asymptotic Localization Fitting", The European Conference on Computer Vision (ECCV), 2018, pp. 618-634

[24] Abhilash Nandy; "A Densenet Based Robust Face Detection Framework", The IEEE International Conference on Computer Vision (ICCV), 2019, pp. 0-0

[25] Jianfeng Wang, Ye Yuan and Gang Yu, "Face Attention Network: An Effective Face Detector for the Occluded Faces", arXiv:1711.07246 [cs.CV], Nov. 2017

[26] R. Wang, J. Tian, and C. Jin, "Joint Face Detection and Alignment Using Focal Loss-Based Multi-task Convolutional Neural Networks," in Biometric Recognition, Springer International Publishing, 2019, pp. 266–273.

[27] X. Zhao, X. Liang, C. Zhao, M. Tang, and J. Wang, "Real-Time Multi-Scale Face Detector on Embedded Devices," Sensors, vol. 19, no. 9, p. 2158, May 2019, doi: 10.3390/s19092158.

[28] X. Xu, H. Duan, Y. Guo, and Y. Deng, "A Cascade Adaboost and CNN Algorithm for Drogue Detection in UAV Autonomous Aerial Refueling," Neurocomputing, Mar. 2020, doi: 10.1016/j.neucom.2019.10.115.

[29] Xin Liu, Yongbin Sun, Ziwei Liu and Dahua Lin, "Learning Diverse Fashion Collocation by Neural Graph Filtering", arXiv:2003.04888 [cs.CV], Mar. 2020

[30] Liaojun Pang, Jiong Chen, Fei Guo, Zhicheng Cao and Heng Zhao, "ROSE: Real One-Stage Effort to Detect the Fingerprint Singular Point Based on Multi-scale Spatial Attention", arXiv:2003.03918 [cs.CV], Mar. 2020

[31] G. Sambasivam and G. D. Opiyo, "A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks," Egyptian Informatics Journal, Mar. 2020, doi: 10.1016/j.eij.2020.02.007.

[32] D. Lee, S. Kim, H. Lee, C. C. Chung, and W.-Y. Kim, "Paved and Unpaved Road Segmentation Using Deep Neural Network," in Communications in Computer and Information Science, Springer Singapore, 2020, pp. 20–28.

[33] Hei Law, Jia Deng; "CornerNet: Detecting Objects as Paired Keypoints", The European Conference on Computer Vision (ECCV), 2018, pp. 734-750

[34]   Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, Nong Sang; " "Learning a Discriminative Feature Network for Semantic Segmentation", The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 1857-1866

[35]   Shaoshuai Shi, Xiaogang Wang, Hongsheng Li; "PointRCNN: 3D Object Proposal Generation and Detection From Point Cloud", The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 770-779

[36]   Dollar, P., Wojek, C., Schiele, B., Perona, P.: Pedestrian detection: An evaluation of the state of the art. IEEE transactions on pattern analysis and machine intelligence 34(4), 743–761 (2012)

[37]   T.-Y. Lin et al., "Microsoft COCO: Common Objects in Context," in Computer Vision – ECCV 2014, Springer International Publishing, 2014, pp. 740–755.

[38]   D. Li, Z. Zhang, X. Chen, H. Ling, K. Huang, A richly annotated dataset for pedestrian attribute recognition, 2016. arXiv: 1603.07054.

[39]   Y. Chenguang, "chinese_ocr," 2018, https://github.com/YCG09/chinese_ocr

[40]   P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabuddhe, and F. Meriaudeau. Indian diabetic retinopathy image dataset (idrid): A database for diabetic retinopathy screening research. Data, 3(3):25, 2018.

[41]   Spanhol F A , Oliveira L S , Petitjean C , et al. A Dataset for Breast Cancer Histopathological Image Classification[J]. IEEE transactions on bio-medical engineering, 2015, 63(7):1455-1462.

[42]   S. G. Armato III, G. McLennan, L. Bidaut et al., "e lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on ct scans," Medical Physics, vol. 38, no. 2, pp. 915–931, 2011.

[43]   Martin Hirzer, Csaba Beleznai, Peter M Roth, and Horst Bischof, "Person re-identification by descriptive and discriminative classification," in Scandinavian conference on Image analysis. Springer, 2011, pp. 91–102.

[44]   Zhang, S., Benenson, R., Schiele, B.: Citypersons: A diverse dataset for pedestrian detection. arXiv preprint arXiv:1702.05693 (2017)

[45]   S. Yang, P. Luo, C. C. Loy, and X. Tang. Wider face: A face detection benchmark. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ◎ (24*7 Support on Whatsapp)