

Review Article

Dense Convolutional Network and Its Application in Medical Image Analysis

Tao Zhou ^{1,2}, XinYu Ye ^{1,2}, HuiLing Lu ³, Xiaomin Zheng,⁴ Shi Qiu,⁵ and YunCan Liu^{1,2}

¹School of Computer Science and Engineering, North Minzu University, Yinchuan 750021, China

²Key Laboratory of Image & Graphics Intelligent Processing of State Ethnic Affairs Commission, North Minzu University, Yinchuan 750021, China

³School of Science, Ningxia Medical University, Yinchuan 750004, China

⁴Research Institute for Reproductive Medicine and Genetic Diseases, Wuxi Maternity and Child Health Hospital, Wuxi, Jiangsu 214002, China

⁵Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China

Correspondence should be addressed to XinYu Ye; 3303626778@qq.com and HuiLing Lu; lu_huilong@163.com

Received 28 February 2022; Accepted 23 March 2022; Published 25 April 2022

Academic Editor: Chen Li

Copyright © 2022 Tao Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Dense convolutional network (DenseNet) is a hot topic in deep learning research in recent years, which has good applications in medical image analysis. In this paper, DenseNet is summarized from the following aspects. First, the basic principle of DenseNet is introduced; second, the development of DenseNet is summarized and analyzed from five aspects: broaden DenseNet structure, lightweight DenseNet structure, dense unit, dense connection mode, and attention mechanism; finally, the application research of DenseNet in the field of medical image analysis is summarized from three aspects: pattern recognition, image segmentation, and object detection. The network structures of DenseNet are systematically summarized in this paper, which has certain positive significance for the research and development of DenseNet.

1. Introduction

Deep learning is an end-to-end approach to extracting and abstracting image features layer by layer and implementing recognition functions. Krizhevsky et al. [1] used deep learning to win the champion of computer vision challenges in ImageNet. Therefore, deep learning methods represented by convolutional neural networks have become a popular research topic in the field of pattern recognition. GoogLeNet, ResNet, and DenseNet have been proposed one after another. Among them, DenseNet [2] has achieved better results among many deep learning models due to the new architecture of dense connectivity, DenseNet achieves interconnection between arbitrary layers and skip connection mode that transfers information from shallow layers directly to deep layers, and enhanced feature transfer and feature reuse between network layers, resulting in a more compact network representation with less feature redundancy. Under

the condition of same layer depth, the convergence performance of network is better, the network degradation and gradient disappearance problems caused by the deepening of convolutional network are better alleviated, and the number of network parameters and computational efficiency are significantly reduced.

Litjens et al. [3] point out that deep learning algorithms, especially convolutional networks, have been successful in medical image analysis. Xie et al. [4] point out that deep learning is widely used in medicine for disease diagnosis, lesion and abnormality detection, lesion and organ segmentation, etc. In recent years, deep learning has developed rapidly in the medical field, represented by companies such as Watson, Microsoft Nuance, Google Health in the world and iFlytek, Tencent Miying, and Ali Health in China. Eagle Pupil Airdoc [5] fundus image-assisted diagnosis system for diabetic retinopathy (DR) has high sensitivity and specificity in detecting, DR. Dr. Wise [6], launched by DeepMed, can

achieve rapid identification, accurate measurement, diagnostic analysis, and scientific follow-up of pneumonia signs, providing clinicians with a basis for treatment, thus realizing precision medicine.

DenseNet has made a breakthrough in medical image analysis tasks [7]; Li and Liu [8] used DenseNet to learn local block features of MRI brain image clusters and achieved a better Alzheimer's disease classification accuracy of 89.5%; Jégou et al. [9] extended densely connected structure to fully convolutional neural network by introducing different depth paths to produce different scale nonlinear mappings for semantic segmentation, thus avoiding gradient disappearance and enabling deeper networks to be trained with fewer parameters; Ke et al. [10] based on 3D DenseNet performed automatic detection and segmentation of nasopharyngeal carcinoma in MRI images, which can obtain higher overall accuracy, sensitivity, and specificity than radiologists. DenseNet can provide clinical aid diagnostic solutions for major diseases such as benign and malignant tumors, cardiovascular and cerebrovascular diseases, and respiratory diseases. Liu et al. [11] based on DenseNet extracted spatiotemporal features of patients' cardiac magnetic resonance imaging (MRI) to predict the ray fraction of left ventricle; Kishan and Janardhan [12] aimed at the problem of small intracerebral hemorrhage and high square difference and used a combination of DenseNet and Inception V3 models to accurately detect and identify early intracranial hemorrhage; Li et al. [13] used transfer learning and DenseNet to classify benign and malignant COVID-19 CT images, reaching the state-of-the-art in terms of accuracy and F1 score.

This paper summarizes DenseNets, and the rest of the structure is as follows: Section 2 introduces the basic principles of DenseNet; Section 3 reviews the development of DenseNet models and summarizes five aspects of broaden network structure, lightweight network structure, dense unit, dense connection mode, and attention mechanism; Section 4 summarizes the application of DenseNet in field of medical image analysis and explores three aspects of pattern recognition, image segmentation, and object detection.

2. Basic Principles of DenseNet

In 2017, Huang et al. [2] proposed DenseNet, a convolutional neural network with densely connected structure, and DenseNet structure is shown in Figure 1.

DenseNet has the characteristics of feature sharing and arbitrary interlayer interconnection. The advantages of DenseNet are feature map reuse through dense connection, reducing interdependence between layers by reusing feature maps from different layers, providing compact and differentiated input features by shortcut connections of different lengths, and effectively reducing the gradient disappearance problem that is difficult to optimize in deep networks. The final prediction is to use features from all layer to obtain better performance and model robustness on a standard dataset with smaller model size and computational effort. The disadvantage of DenseNet is that the feature maps of each layer are spliced with the previous layer, and the data is replicated multiple times. As the number of network layers increases,

the number of model parameters grows linearly, eventually leading to explosive growth in computation and memory overhead during training.

3. Development of DenseNet

There are five improvement methods of DenseNet, such as broaden DenseNet structure, lightweight DenseNet structure, dense unit, dense connection mode, and attention mechanism. This section summarizes the development of DenseNet, as shown in Figure 2.

In 1998, LeCun et al. [14] proposed LeNet model, in which some new operators such as convolutional operator, pooling operator, and full connection operator were used, and LeNet become a pioneer model in deep learning models.

In 2012, Krizhevsky et al. [1] proposed AlexNet, in which ReLU was used for activation function; network architecture uses dual GPUs, reduced pooling steps, enhanced training data, and dropout function to avoid overfitting; this network achieved 10.9% better result than second place in ILSVRC competition.

In 2015, He et al. [15] proposed ResNet for the problem of network with more stacked structural layers but fast performance degradation, using residual connections to transfer shallow features directly to deeper layers, so that the originally fitted output $H(x)$ becomes the residual output $H(x) - x$. Even when learning structural layer with 0 features, ResNet only does constant mapping and learns new features based on input features; ResNet learns new features based on input features, alleviating the degradation problem that occurs in deep models and giving network better performance.

In 2017, Huang et al. [2] further improved feature reuse capability based on ResNet and proposed DenseNet network with dense connection operation, but in the case of dense connection as the number of layers of dense blocks increases, input feature map dimension at the proximal end becomes larger and larger, to alleviate this problem Huang et al. [2] added 1×1 convolution (bottleneck unit) into dense unit to reduce the dimension of input feature maps, reducing the computational effort while incorporating different dense unit features. DenseNet-B uses 1×1 convolution to reduce input feature map dimension of each layer to a fixed multiple of growth rate; DenseNet-C adds compression rate in transition layer to reduce output feature dimension of dense blocks; DenseNet-BC adds both bottleneck unit and variable compression rate transition layer.

Broaden DenseNet structure is multibranch improvement of network single layer, network structure block, and whole network using multibranch extraction function, including multipath structure, parallel dense block, parallel transition layer, and parallel network.

Lightweight DenseNet structure significantly reduces the number of network parameters and computation from two aspects: network model architecture design and network model compression. Network model architecture design which includes grouped convolution, depth-separable convolution, simplified dense block, one-time aggregation, and network model compression has automatic sparseness.

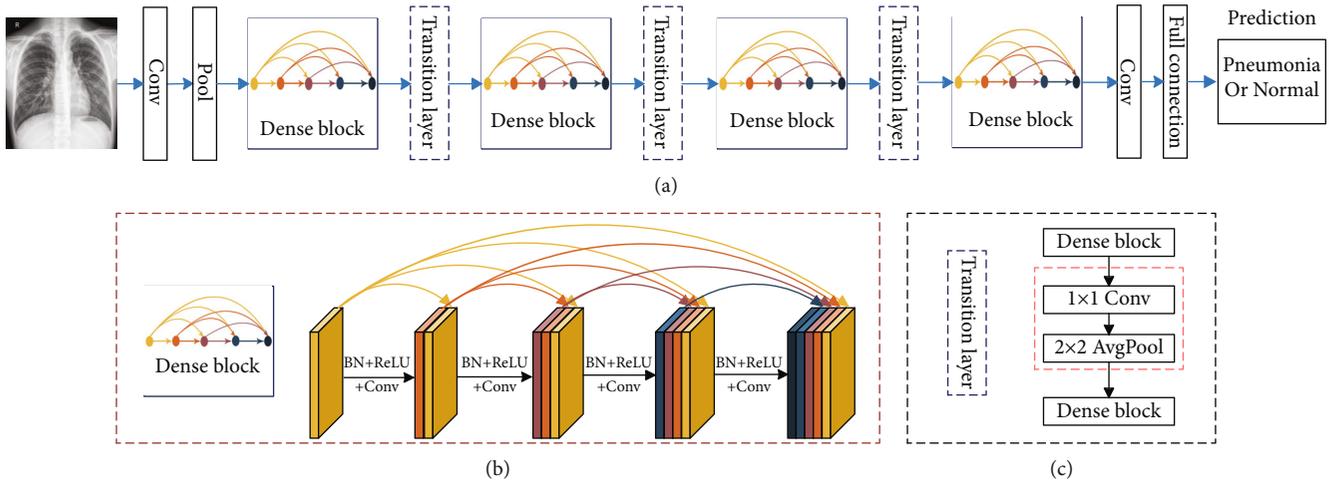


FIGURE 1: DenseNet structure. (a) The basic structure of DenseNet consists of dense block, transition layer, convolutional layer, and fully connected layer. (b) Denseblock consists of densely connected dense units with nonlinear mapping functions of BN, ReLU, and Conv, which are designed with preactivation strategy to make network training easier and generalization performance better. Dense unit input is spliced and merged with all outputs of the previous dense units, and new features generated also need to be passed to subsequent dense units, so that shallow features of dense block are repeatedly reused and effectively utilized, which can alleviate gradient disappearance to a certain extent, and a large number of features can be generated with a small number of convolution kernels; final DenseNet model is relatively in scale. (c) Transition layer is the structure between adjacent dense blocks, which consists of 1×1 convolution and 2×2 average pooling layer, compressing dense block input and all extracted feature information, reducing feature map size and dimensionality, which can effectively reduce the number of dense block parameters and prevent network from overfitting. The fully connected layer is classification prediction layer, which reduces the influence of feature location on classification by integrating category feature information in network features, and classifies feature information after weighting.

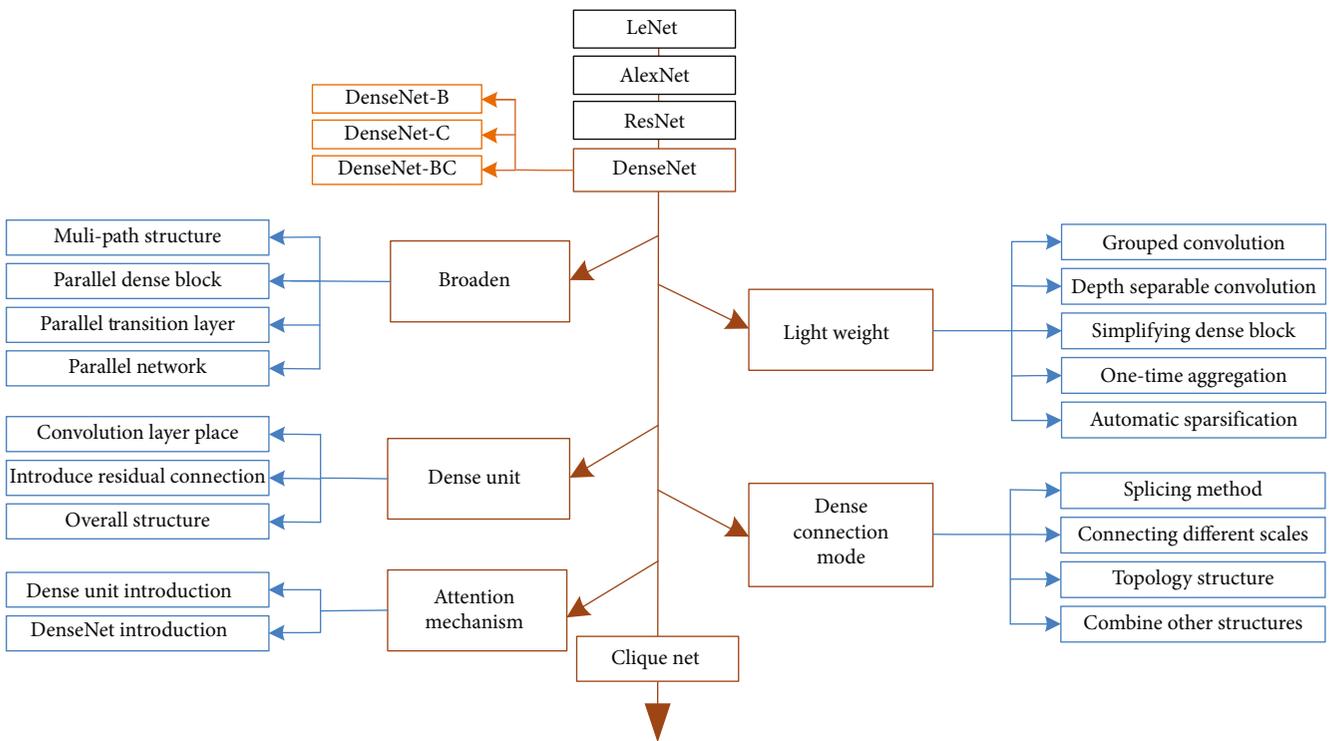


FIGURE 2: Development of DenseNet.

Dense unit is composite function improvement of DenseNet, optimizing feature extraction module dense unit for specific applications, summarized in three aspects: convolution layer placement, introduce residual connection, and overall structure.

Dense connection mode is a study of the way dense blocks are composed, dense connection improvement has four aspects: splicing method, connect dense unit of different scales, topological structure, and combine with other structure.

Attention mechanism is to emphasize or select important features in DenseNet and suppress some irrelevant detailed information, which can be divided into attention mechanism in dense unit and DenseNet.

Chen et al. [16] indicated that densely connected path of DenseNet increases linearly with increasing network depth, and the parameters increase dramatically; for this reason, Yang et al. [17] proposed CliqueNet, a convolutional neural network with alternating update mechanism, Clique block performs same dense propagation as DenseNet, and then, all layers except the top layer are used as the bottom layer for alternating updates, stitching feature maps of each layer into the output of Clique block, passing only fixed-dimensional feature maps, with forward and backward connections between any layer, cycling, and alternating updates to maximize the flow of information between layers. Final network uses fewer convolutional kernels to generate a large number of features, which greatly improves the parameter efficiency, and parameter propagation process generates multilevel features by cycling; experiments show that the second stage of Clique block helps to suppress noise and improve performance.

3.1. Broaden DenseNet Structure. Depth is not necessarily good, and width is not necessarily bad. On the one hand, in 2014, Ba and Caruana [18] pointed out that a well-trained deep network can usually find a suitable shallow network instead, and in 2019, Hanin and Rolnick [19] pointed out that expressiveness of network in practical applications is independent of network depth and only grows linearly with the number of neurons; on the other hand, deep learning models currently are trained by GPUs. GPU parallel processing makes it easier to train widened networks. Therefore, depth and width are not completely opposed to each other, increasing both depth and width is increasing the number of learnable parameters and improving network fitting ability, increasing depth can obtain a larger sensory field and capture more pixel-like features, and increasing width can obtain more subtle and richer features.

Broaden DenseNet structure is using multiple branches to improve network single layer, structure block, and whole, so that network learns more rich features per layer, such as texture features with different directions and frequencies. Although network deepening can extract knowledge layer by layer abstraction and continuous refinement, too narrow network can capture limited patterns per layer, and the network is no deeper to extract information. Broaden DenseNet structure approach is shown in Figure 3; there are multipath

structure, parallel dense block, parallel transition layer, and parallel network.

First, multipath structure is an improvement of multipath feature extraction layer on single-path feature extraction, using multiple scales to achieve different perceptual fields, increasing the adaptability of network to features of different sizes and enriching features. Multipath convolution also uses the computational principle of decomposing sparse matrices into dense matrices to speed up convergence, aggregating output features with strong relevance, decomposing feature sets into multiple densely distributed subsets, aggregating features with strong relevance, and finally outputting less redundant information and faster convergence. In 2014, Google's Szegedy et al. [20] proposed typical multipath structure inception, which is parallel to commonly used 1×1 , 3×3 , 5×5 convolutions and 3×3 max-pooling to achieve different perceptual fields for extracting rich features and making classification judgments more accurate. Li et al. [21] proposed DenseNet-II for benign and malignant classification of mammogram images; the inception structure was used to improve adaptability of network to different scales and maintain the sparsity of network structure, while efficiently expanding network depth and width, resulting in better classification performance with good generalization and robustness. Arega and Bricq [22] used cascaded segmentation network FC-DenseNet for automatic segmentation of cardiac scars in multisequence cardiac MRI; similar inception structure captures relevant features such as scars and edema of different sizes, significantly improving segmentation results with the same computational overhead.

Second, parallel dense block is original single-branch dense block multiple times in parallel to increase network width, combining the benefits of network depth and width to improve model robustness; although deep networks have higher performance, they are limited in terms of gradient disappearance, reduced forward flow, and slower training time. Lodhi and Kang [23] proposed multipath-DenseNet with multiple dense blocks parallel learning to find valid paths from some or all of the blocks, which allows efficient gradient flow and achieves higher accuracy than other baselines using fewer parameters on the classification dataset. Chen et al. [24] proposed MFR-DenseNet with multipath feature recalibration, where SE module assembles multipath dense blocks from different channel feature maps to build interchannel feature correlation model to achieve strong feature extraction, which achieves high accuracy in classification while maintaining training stability.

Third, parallel transition layer is a branching supplement based on the single-branch transition layer to realize broadening. Liu et al. [25] proposed dense binary tree DenseBT-Net for lung nodule classification, which introduces central cropping operation into transition layer and parallelizes it to create indirect shortcuts between different dense blocks, preserving detail features of nodule morphology and filtering irrelevant background information; binary tree-like singular network model can generate rich multiscale features, making the parameter scale lighter and with higher parameter efficiency.

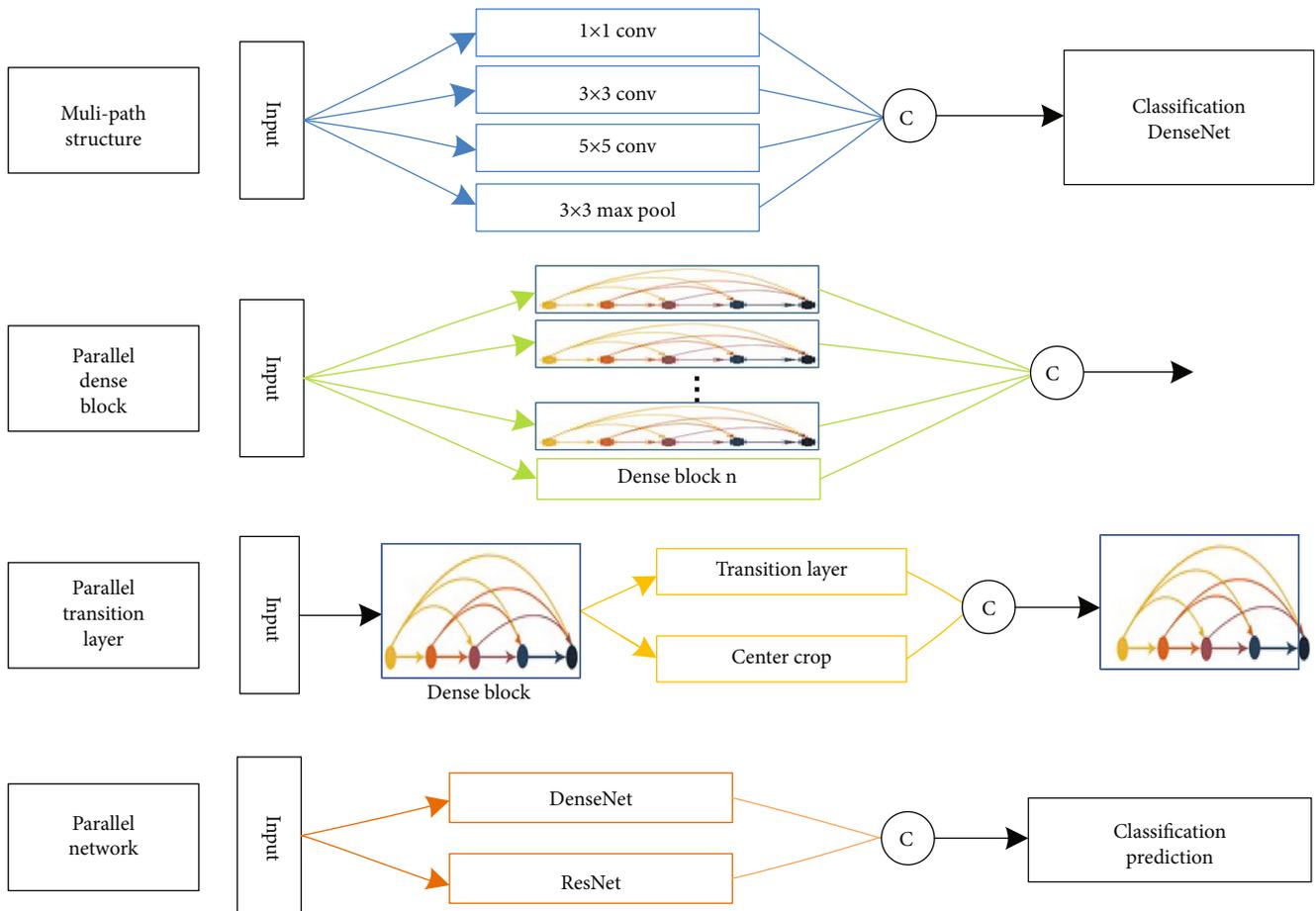


FIGURE 3: Broaden DenseNet structure.

Fourth, parallel network is multiple networks in parallel, and the fusion of multiple network features can ensure more stable and accurate prediction. Chen et al. [16] argued that ResNet is essentially DenseNet with shared connections, and ResNet implicitly has residual path reuse features, and dense connections that can be mined for new features, based on this joint perspective dual-channel network (DPN) were proposed, by sharing same features of network through dual-path structure; high accuracy is achieved with flexibility, small models, less computation, and low resources. Nguyen et al. [26] proposed MSDENSE-DAT to predict regions in crowd scenes that attract human attention, designed two-branch homography network based on DenseNet201 to extract multiscale features, extracted as many features as possible from original size and half-size images, and cascaded two-branch features along channel axis, and also designed self-attention blocks to emphasize interfeature correlation; MSDENSE-DAT extracts the best features with low density in the crowd.

3.2. Lightweight DenseNet Structure. Lightweight of CNN is to reduce the number of parameters and computation without degrading the performance of existing models and to achieve storage and computation on mobile platforms with limited storage space and power consumption. CNN is

widely used in image analysis and has achieved great success, greatly reducing the number of parameters and computation through weight sharing and sparse connection; however, two aspects, model storage and model prediction speed, limit the application of CNN in mobile. Lightweight of CNN is important research in deep learning, and two main directions are studied: lightweight model architecture design and model compression. This section summarizes four improvements of lightweight DenseNet structure in model architecture design: grouped convolution, depth separable convolution, simplified dense block and one-time aggregation, and automatic sparseness of model compression, as shown in Figure 4.

First, grouped convolution is to group feature maps and then convolve them separately, the number of model parameters decreases as the number of groups decrease, and model structure and training are more efficient. In 2012, AlexNet [1] proposed grouped convolution and utilized dual GPU processing. To apply DenseNet to mobile devices with limited computational resources, Huang et al. [27] proposed CondenseNet based on grouped convolution, where the network is connected layer by layer, preserving dense connectivity of DenseNet, achieving feature reuse, sharing input features using self-learning grouped convolution modules, eliminating layer-to-layer connections, and avoiding

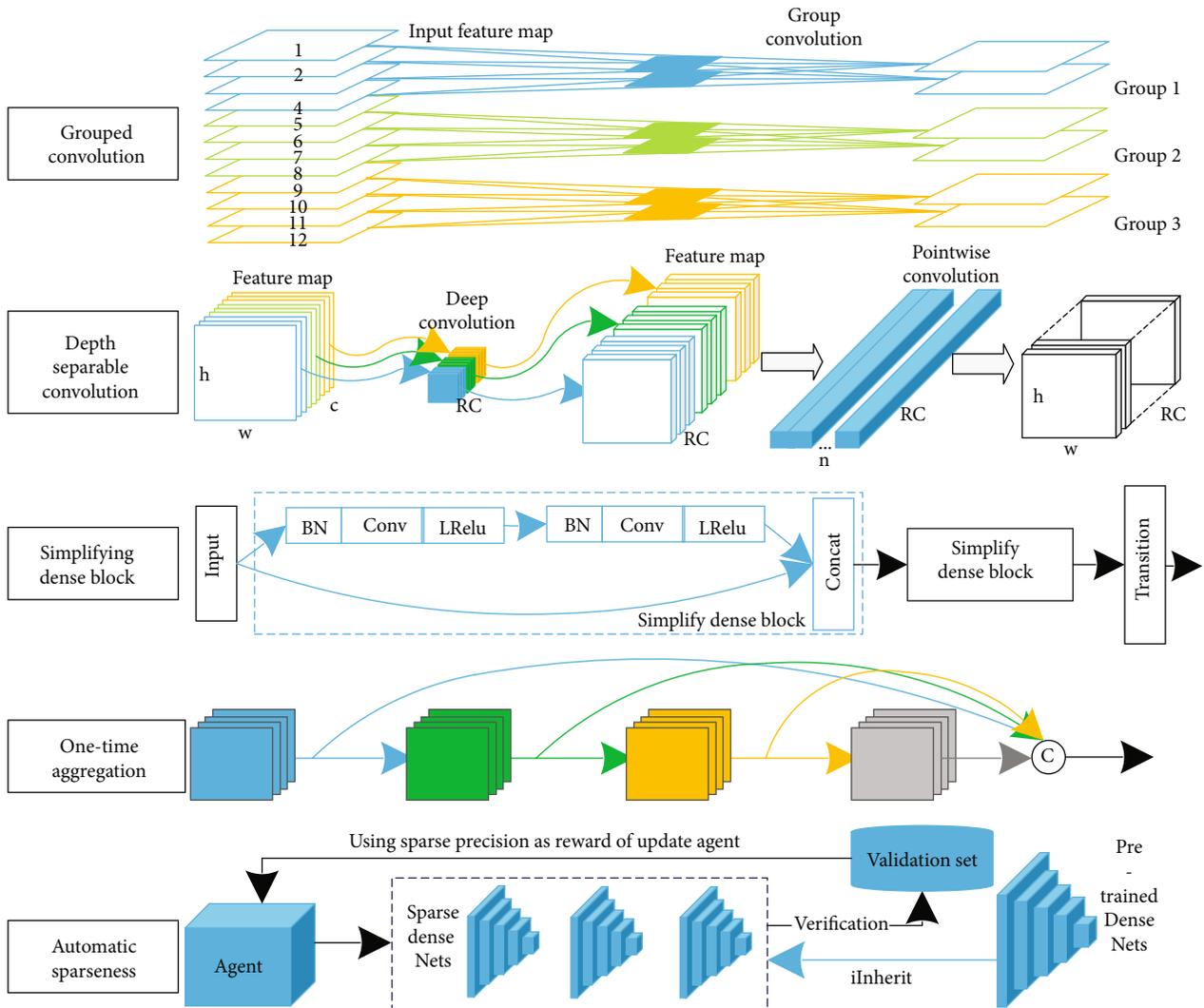


FIGURE 4: Lightweight DenseNet structure.

accuracy branching process that affects too much; CondenseNet uses only 1/10th of computation at an accuracy level comparable to that of DenseNet. Li et al. [28] proposed Dense2Net with efficient multilevel group convolution, using channel mixing to rearrange features, tiling features into multiple groups and randomly receiving the previous group of features, while fusing all groups of different scales into output using 1×1 convolution to improve parameter efficiency and information transfer; Dense2Net has higher accuracy and better information transfer than DenseNet.

Second, depth separable convolution (DWC) is a decomposition of the convolution strategy into spatial convolution followed by channel convolution, replacing convolution layer to achieve a lightweight network, using a filter for each input channel in spatial directional convolution (DWC), and applying 1×1 point-by-point convolution (PWC) on the channel side to linearly combine the depth directional convolution outputs. MobileNet [29] uses depth separable convolution to build a lightweight deep network

that greatly reduces the computational effort with a small reduction in accuracy, achieves a balance between model efficiency and classification accuracy, and performs well in task such as target detection, fine classification, and face attribute. Chen et al. [30] proposed Mobile-DANet, an attention-embedded lightweight network for corn disease recognition, using depth-separable convolution instead of convolutional layers in dense blocks and depth convolution using input mapping filters to achieve channel-by-channel convolution and combining channel-by-channel results through 1×1 convolution, while using spatial and channel attention mechanism to detect and highlight targets in multifeature map, with the final model featuring small size, high accuracy, and excellent performance.

Third, simplified dense block is to replace original dense block with multiple simplified dense blocks in series, fixing the number of dense units to be smaller and multiple simplified dense blocks in series to replace original dense blocks, reducing output feature map dimension while maintaining

feature reuse. s-DenseNet was proposed by Yu et al. [31] for pathological grading of breast cancer on large-size mammograms, using fewer dense units to prevent network overfitting, extracting high-level semantic features, and combining logistic regression to select image histology feature values, the trained network achieves high accuracy at lower complexity. Huang et al. [32] designed a lightweight neural network using simplified dense block consisting of two dense units, stitching adjacent blocks of mammogram and osteosarcoma histological images as input; feature vectors were mapped to simplified dense block for kernel learning; final network was easier to train and better in terms of classification accuracy, sensitivity, and specificity.

Fourth, one-time aggregation is to aggregate outputs of all dense units to the end of dense block, rather than two-by-two interconnection of dense units. Lightweight network design considers two main factors, model size and FLOPs, to reduce inference time and energy consumption; memory access cost (MAC) and GPU computational efficiency are also considered. Densely connected aggregations of intermediate features with different sensory fields can produce powerful features using a small number of parameters and triggers, but dense connections can lead to a large memory access overheads, and 1×1 convolution using a small number of operations can also lead to inefficient GPU computation. To address the slower and inefficient DenseNet detector, a fast and efficient architecture VoVNet consisting of one-time aggregation (OSA) was proposed by Lee et al. [33], which reduces memory access cost and performs efficient computation by using diverse feature representations with multiple reception rates; dense aggregation is aggregating all previous features to subsequent layers; while OSA is aggregating all features at once to final unit, it also removes bottleneck unit and keeps intermediate layer input constant, greatly improving MAC and GPU efficiency while maintaining cascade strength; VoVNet applied to target detector achieves twice detection speed, and energy consumption is $1.6 \times -4.1 \times$ lower than DenseNet.

Fifth, automatic sparsification method for model compression to achieve lightweight is to prune DenseNet by reinforcement learning. Li et al. [34] proposed automatic sparsification (ADS) method to prune redundant jump connections in DenseNet and ADS compression based on reinforcement learning (RL) method to prune dense connections representation by the original adjacency matrix; sparse DenseNet verification accuracy is used as a reward and punishment for updating RL agents, and high-performance sparse networks are generated by iterating RL agents with fewer parameters, less computation, and portability after sparse.

3.3. Dense Unit. DenseNet uses dense cells interconnected two by two to form dense blocks; this section summarizes three aspects of convolution layer placement, and introduces residual connection and overall structure.

3.3.1. Convolution Layer Placement. Dense unit of DenseNet uses preactivation structure, in which BN and ReLU layers achieved preactivation before convolution, and also uses the structure of 1×1 convolution followed by 3×3 convolu-

tion. Dense unit is improved by placement of convolution layer, which can be divided into preactivation and nonpreactivation, where ReLU and BN can interchange their positions, and the specific improved form is shown in Figure 5.

First, convolutional layer is easier to train and has better generalization performance after BN algorithm and preactivation structure of activation function. Zhu et al. [35] achieved image superresolution using dense jump-connected network, where dense unit provides large sensory fields using 3×3 convolution, and then, 1×1 convolution deepens network to learn robust feature representations, and small filters with a small number of parameters are used to deepen network, making full use of hierarchical feature deepening structure, and network outperforms new algorithm with robust fitting capability. Wei and Liu [36] introduced dense blocks in SSD network to detect hazardous substances in X-ray images. Dense unit connects two 1×1 convolutions in series to avoid destruction of learned feature map region information, and the method has good transfer learning capability.

Second, convolutional layer does not use preactivated dense units before BN and activation function. Wei et al. [37] added dense blocks of first convolution in YoloV3 backbone, and transition layer also used dense units in parallel with maximum pooling layer, and the model was effective in multiobjective defect detection of railroad routes. Lei et al. [38] used densely connected dilated convolution for codec connection, and dense blocks were used for first convolution of dilated dense unit, which can enhance effective feature transfer and retain finer structural information for accurate segmentation of skin lesions.

Third, convolutional layer distributes BN-normalized data to the nonsaturated region before activation function and BN, and ReLU can then control the degree of saturation of activation; if the order of BN and ReLU is switched, ReLU will lead to deactivation of some of BN neuron, which causes BN instability and affects the model performance, practically different application scenarios have different role in the order. Hasan et al. [39] proposed an automatic semantic segmentation network (DSnet) for skin lesion segmentation, in which the encoder uses dense blocks and transition layers to avoid learning redundant features and uses BN of dense units after activation function, and DSnet outperforms existing baseline segmentation networks in several metrics.

Fourth, convolutional layer precedes activation function, and BN is more suitable for denoising, but it leads to slow convergence in superresolution and large fluctuation in the loss function, which also increases the memory and computational burden. Mustafa et al. [40] proposed multilevel dense MLDNet for multifocus image fusion (MFIF), which uses parallel dense blocks to extract multilevel local dense and global features in the source image; dense units are connected using convolution followed by ReLU to retain as many dense features as possible, and MLDNet has significant performance improvement over the latest methods. Li et al. [41] proposed multiresolution dense network to connect mutually different structures dense units densely, with different size convolution for feature extraction of three central consistent lung field blocks with different resolutions,

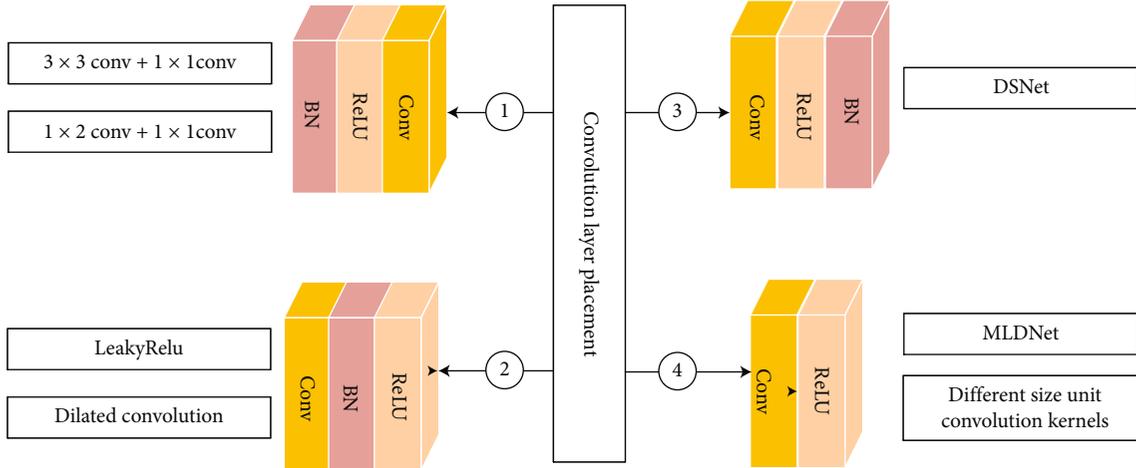


FIGURE 5: Convolution layer placement.

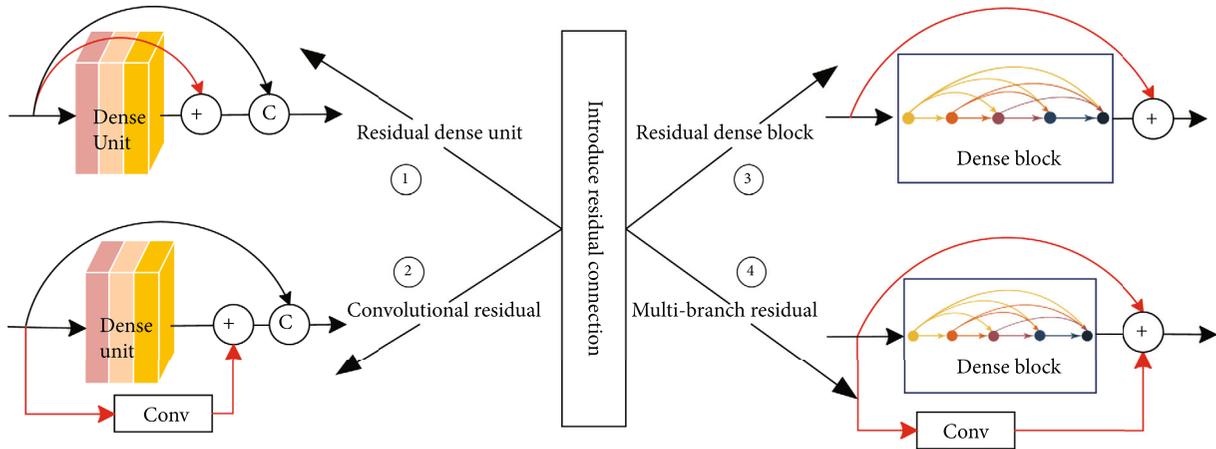


FIGURE 6: Introduce residual connection.

and fusion to detect lung nodules, which outperformed general JSRT database for radiologists by up to 99%.

3.3.2. Introduce Residual Connection. Residual connection uses skip connection to connect shallow layers of information directly to deep layers, and output is expressed as a linear superposition of input and its nonlinear transformations, using a new expression to achieve a constant mapping between layers. Deep learning relies on error chain derivative back propagation for parameter update; too small derivatives will become smaller and smaller gradients after multiple series multiplications, leading to gradient dispersion; residual connection adds constant terms to derivatives, so that error signal is propagated directly to the shallow layer without intermediate weight matrix changes, which alleviates the problem of gradient dispersion to a certain extent, and information is propagated back and forth more smoothly. Introduction of residual connection in DenseNet can solve network degradation problem, thus deepening network model and improving performance. Four structures are introduced for residual connectivity: residual dense unit,

convolutional residual dense unit, residual dense block, and multibranch residual dense block, as shown in Figure 6.

First, residual dense unit is residual connection between inputs and outputs of dense unit, Feng et al. [42] proposed recurrent network DCRN to cope with single-image super-resolution (SISR) task, which performs local residual learning within dense block and dense unit to extract semantic features, and experiments on Set5, Set14, BSD100, and Urban100 datasets to demonstrate DCRN efficiency and model component effectiveness.

Second, convolutional residual dense unit is a residual connection using convolutional improvements to prevent gradient disappearance or explosion problems while speeding up convergence. Yang et al. [43] introduced convolutional residual in DenseNet by parallelizing multiple multiresolution convolutional residuals in dense block, effectively reducing the number of superresolution network parameters, and increasing model hierarchy and network depth, with prediction speed reaches 25 ms per image under certain accuracy loss. Single-branch convolutional residual block can only extract single-level semantic information;

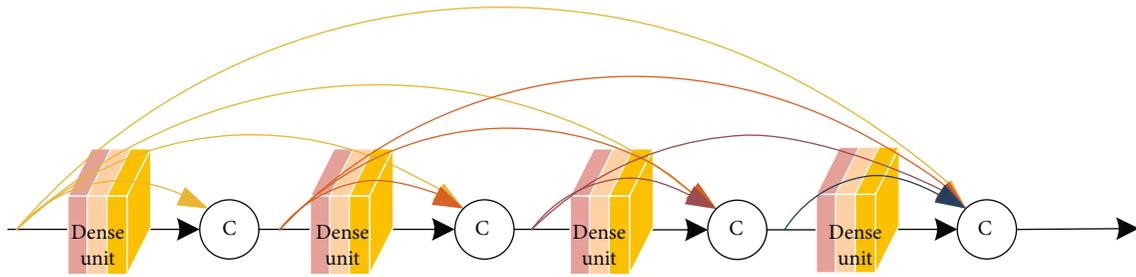


FIGURE 7: Dense connection mode.

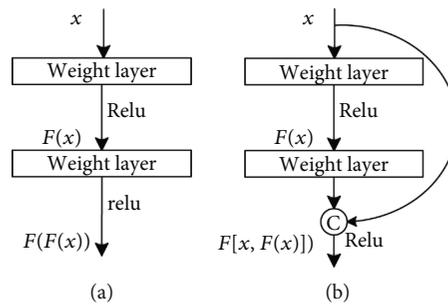


FIGURE 8: (a) Traditional CNN unit mode and (b) DenseNet unit mode. (a) Input gradient information in conventional unit module goes through many layers. (b) DenseNet unit module uses dense jump connection to splice input features with learned features as subsequent layer inputs thus achieving feature reuse, which is easy to train and more parameter efficient in DenseNet. Specific structural design details can be expressed in following equation with some notations defined in advance: input picture is X_0 , l is index of neural network layers, $H(\bullet)$ indicates nonlinear operations combination, and $H(\bullet)$ in DenseNet refers to composite function consisting of BN, ReLU, and Conv.

Wang et al. [44] designed multiscale residual network (EMRN) in image superresolution task and densely connected multiscale residual blocks extracted image hierarchical features to achieve multilevel semantic information in different perceptual domains.

Third, residual dense block is dense block using residual connection between inputs and outputs, and residual dense block inherits advantages of ResNet and DenseNet, improves information flow of network, and retains more structural information. Wu et al. [45] proposed dense pyramid residual network (DPRnet) for restoring fog-free images, residual dense block extracts network features and quadruple downsampling features, global residual connection is applied to whole network, allowing the feature maps to be heavily reused and deepening network at different scales, and ultimately, DPRnet has better performance.

Fourth, multibranch residual dense block is dense block that uses both residual connection and convolutional residual connection, and multibranch residual structure improves diversity of feature extraction for hierarchical feature extraction. Long et al. [46] proposed aggregated residual dense network (RXDNFuse), which fuses residuals and convolutional residuals from infrared and visible images into parallel dense blocks to extract hierarchical features, and qualitative evaluation showed that rich texture details and prominent thermal radiation information were effectively retained.

3.3.3. Overall Structure of Dense Unit. DenseNet is connected with dense units in a dense connection, and overall

structure of dense unit is replaced with whole structure; five overall structure improvements are as follows.

First, multirate extended convolutional structure MDSC replaces dense unit, 1×1 convolution, and four different extended rate 3×3 convolutions to perform input feature mapping in parallel, which was used in training DenseNet by Zhang et al. [47], using original image and marker point set, adding MDSC to increase network width laterally to fully utilize multiscale vascular spatial feature information at lower computational cost, and dense connection alleviates gradient disappearance and accelerates network, the average dice score reached 93.20% with good integrity and sensitivity for cerebral vessel segmentation, especially for slender vessels. Second, Wang et al. [48] proposed dual residual attention model (DRAM) for SISR task, in which dense units tandem two 3×3 convolutions capture more global information, channel attention features and convolutional layer features share information in tandem, spatial attention features and convolutional path features are fused, and dual feature fusion enables the model to capture important high-frequency components; the superiority of model in benchmark dataset was verified. Third, Zeng et al. [49] used densely connected cascade network DCCN for MRI reconstruction of K -space undersampled data, using dense unit tandem residual dense block (RDB) and data consistency layer (DC), with RDB combining dense connection and residual learning to fully utilize features of different layers, and DC used data consistency to improve MRI image reconstruction under different sampling trajectories and sampling

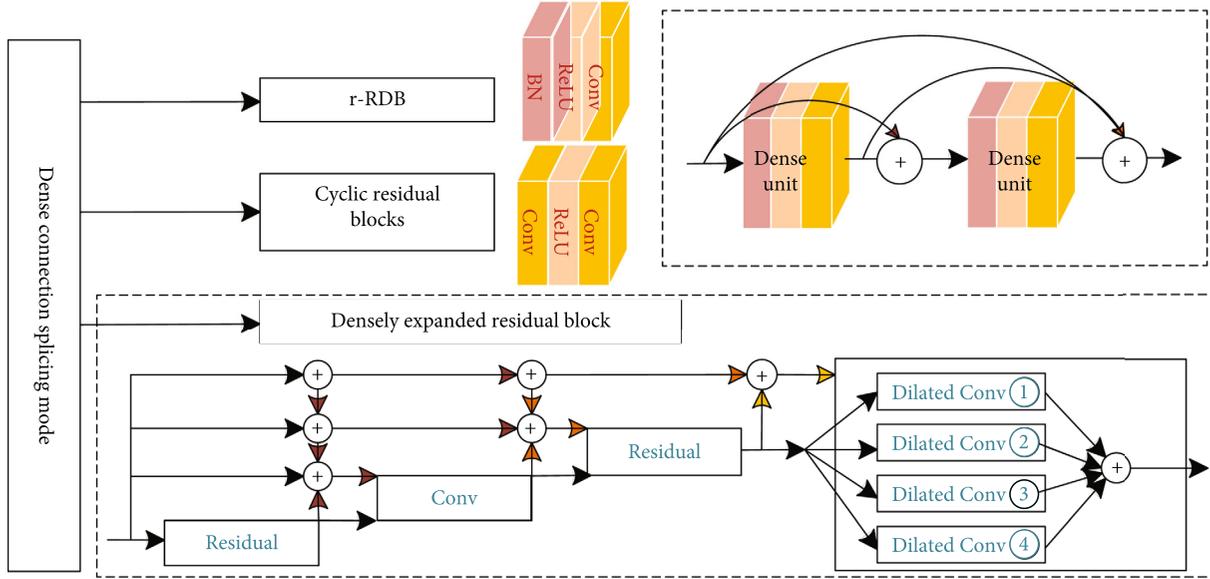


FIGURE 9: Dense connection splicing mode.

rates. Fourth, Zhang et al. [50] used densely connected inception-Res module for segmentation of medical images, with branches of different kernel sizes aggregating feature maps, widening network to learn more features, and residual connection referencing input feature mapping learning function to make the module easier to learn; final model achieved better performance on lung and brain tumor segmentation datasets. Fifth, Qamar et al. [51] proposed encoder dense block with residual connection for skin lesion segmentation architecture, using residual connection to obtain more robust features, dense unit also uses expanded space pyramid pools (ASPP) to efficiently capture global multiscale contextual information while maintaining network parameters, ASPP also aggregates image space information, and the results of study showed that architecture achieved the most advanced performance.

3.4. Dense Connection Mode. Dense connection mode is two-by-two interconnection of all dense units in same dense block, and the feature maps of same scale are connected using channel stitching, as shown in Figure 7.

Dense connection mode [9] maximums information flow and reduces the parameters, while alleviating network gradient disappearance and overfitting phenomena, dense connection has following features: first, dense connection uses output as input to subsequent layers, enhancing feature transfer and using features more efficiently; second, dense unit in dense connection is narrower, adding only a small number of feature maps to network each time without relearning redundant feature maps, and final classifier makes predictions based on all features, maintaining network performance while reducing the number of parameters; third, each layer in dense connection directly uses gradient of loss function and initial input information, which is equivalent to invisible depth supervision and helps alleviate gradient disappearance; fourth, dense connection also has a regularization effect, which has a suppressive effect on overfitting,

and network also can mitigate overfitting with fewer parameters. Dense connection improvement has four aspects: splicing method, connect dense unit of different scales, topological structure, and combine with other structure.

3.4.1. Splicing Method. Current unit input in dense connection is stitched from previous unit feature maps, and stitching method is to perform channel stitching at same feature map size. Traditional convolutional neural network unit module is different from dense connection neural network unit module, as shown in Figure 8.

Output of the first layer of conventional neural network can be expressed as

$$X_l = H(X_{l-1}). \quad (1)$$

DenseNet splices different layer features, output of the first layer is expressed as:

$$X_l = H([X_0, X_1, \dots, X_{l-1}]), \quad (2)$$

where $[X_0, X_1, \dots, X_{l-1}]$ represents $0 \sim (l-1)$ layer feature map splicing; dense connection mode is shown in Figure 9.

First is residual dense connection mode for channel fusion. Zhang et al. [52] proposed improved residual dense blocks (r-RDB) in image string and handwritten character recognition tasks, maintaining RDB local feature fusion and local residual learning capabilities, cascading multiple r-RDBs to form global dense block using dense connection of residual summation, and adaptively learning global dense residual features to ensure interblock information flow, r-RDB feature fusion, and residual learning refine block structure and reduce inner computational cost. Adegun and Viriri [53] proposed DenseNet framework for cyclic residual block classification based on dense block, where dense connection within cyclic residual block is combined using

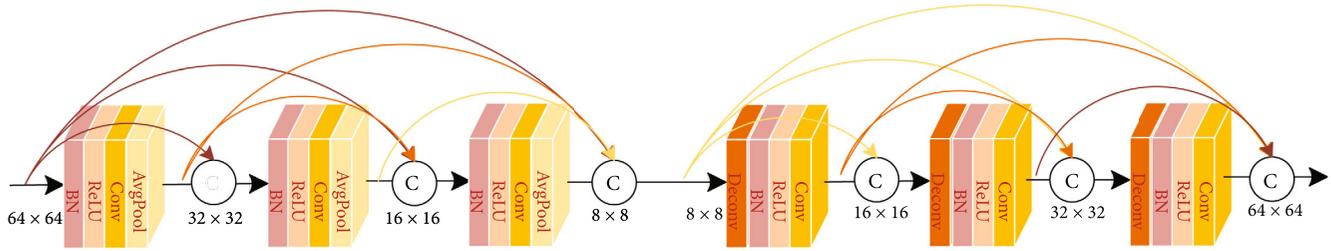


FIGURE 10: Connect dense units of different scales [55].

summation operations to reduce computational effort, contour refinement, and localization of lesion boundaries for edge potential; results showed that effective preprocessing and segmentation of skin lesions can improve classification performance.

Second, dense-expanded residual block improves dense connectivity as a method for multiplexed residual fusion, Hong et al. [54] proposed an end-to-end network MMCL-Net for simultaneous detection, segmentation, and classification of spinal structures, establishing spatial dynamic connectivity for three structures and three tasks, dense connectivity for multilevel aggregation of expanded residual unit to extract important radiological features, and dense connectivity as residual mapping; shared learning among multiple tasks represents adaptive optimization of model, and densely expanded residual module outputs are received in parallel by dilated convolutional layers with different expansion rates; MMCL has been validated on MR images of more than 200 clinical patients, and it can accurately identify structurally relevant lesions in spine and show superiority in grading lumbar spinal neural foraminal stenosis.

3.4.2. Connect Dense Units of Different Scales. Dense connection in DenseNet is to connect same scale dense unit feature maps, and connecting dense units of different scale is shown in Figure 10.

Dense connection is used between different scale structures of encoders or decoders. Wang et al. [56] used dense connection within encoder and decoder networks, average pooling between convolutional layers to eliminate redundant features, and dense connectivity of units with different feature sizes to adequately extract biomarkers associated with recurrence of high-grade plasmacytic ovarian cancer from CT data; the model provides a new prognostic analysis approach. Pemasiri et al. [57] used multimodal semantic segmentation model to segment human body parts in visible and X-rays, where densely connected features were adjusted to dense units of same size, and features of different modalities were integrated using channel splicing to achieve single model for multimodal semantic segmentation. There is only one set of cascade layers between U-Net codec blocks, and Jun et al. [58] designed nested codec architecture (T-Net) with pooling and upsampling inserted in codec blocks, feature mappings interconnected, encoder low-level and high-level features connected to whole decoder, and feature mappings at different scales to produce complex semantic segmentation, and experiments showed that Dice similarity

coefficient of T-Net reached 83.77%, which is 10.69% higher than that of U-Net.

3.4.3. Topological Structure. Conventional L -layer convolutional neural network has L connections, while DenseNet has $L(L+1)/2$ connections, densely connected L th layer has L inputs, which are feature maps of its predecessors, and its feature maps also need to be passed to all subsequent layers, as shown in Figure 11. Due to this two-by-two interconnected connection, DenseNet has constant mapping, deep supervision, and deep diversity characteristics, densely connected approach has a simple internal representation and can reduce redundant features.

First, dense connection topological structure is to aggregate outputs of all dense unit to the end of dense block. In superresolution (SR) domain, deeper network helps to improve image reconstruction quality and introduce residual connection to avoid gradient disappearance, but deep network is computationally intensive and inefficient, Yang et al. [59] used a new dense connection approach to extract low-resolution image features; model topology is to merge outputs of unit into the end of module and obtain deeper network and richer feature map by reducing the number of channels and network parameters; model was experimentally shown to be effective in SR with different magnifications. Mostefa et al. [60] proposed DenseMultiOCM for extracting MRI image features in brain tumor segmentation, improved dense connection by connecting only unit output to final output of dense block, dense units in series to two 3×3 convolutions, and final unit used pooling and upsampling to extract half-scale features; experiments showed model segmentation results were improved.

Second, residual connection is introduced in dense connection mode, two methods of SR are to learn features in high-resolution (HR) and low-resolution (LR) image space, and HR method is to interpolate LR information into HR, but it increases computational complexity and may introduce additional noise; LR method is to extract features and then increase spatial resolution by deconvolution, which is less computationally intensive, but difficult to learn multiscale nonlinear LR-HR mapping. Both methods learn mapping relationships in single-scale image space and cannot provide SR information across different scale of features, and although there is a Laplace pyramid structure to progressively upsample LR information to generate multiscale features, it also uses only highest-level features and ignore low-level features, Zhou et al. [61] designed dense convolutional autoencoding

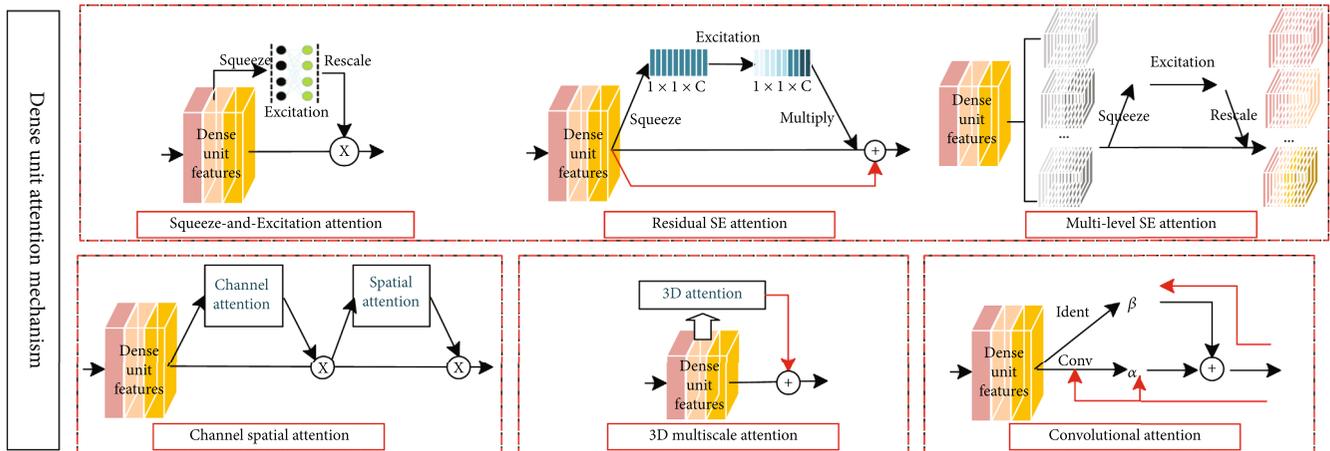


FIGURE 12: Attention mechanism in dense unit.

3.5. *Attention Mechanism.* Attention mechanism [68] is introduced into codec framework, which makes it easier to learn the intermapping relationships between multiple contents and different modal data; thus, allowing better representation and overcoming its uninterpretable and hard to design drawbacks, attention mechanism does not require supervised signals and appears to be extremely effective for problems with little a priori cognition. Attention module reduces computational burden of high-dimensional data by structurally selecting a subset of input and improves the quality of output by “eliminate false and retain true,” allowing model to focus more on useful information in input that is significantly relevant to current output. Introduce attention mechanism into dense unit, and DenseNet is based on three main considerations: first, allow model to achieve better performance in text image analysis tasks; second, attention mechanism can calculate weights for each input and improve the interpretability of network model operates way [69]; third, it can alleviate some defects in recurrent neural networks, such as performance degradation due to increasing sequence input length and computational inefficiency due to input ordering.

3.5.1. *Attention Mechanism in Dense Unit.* Dense unit introduces attention mechanisms and mainly squeezes excitation (SE), residual SE, multilevel SE, channel spatial, 3D multiscale, and convolutional attention mechanism, as shown in Figure 12.

First, SE attention mechanism is to increase skip connection of attention mechanism in adjacent layers, focuses on interchannel relationships and learns the importance of different channel features, and improves interchannel information interaction, thus improving network performance. Feng et al. [42] proposed densely connected recurrent network (DCRN) for single-image superresolution tasks, introducing SE for dense unit and dense block features, rescaling and extracting contextual information, and global pooling of SE attention mechanism to extract global structural information, and fully connected layer emphasizes information channels and suppresses useless channels to improve model representation and reconstructed image consistency. Huang

et al. [55] proposed a preprocessing method to optimize diagnostic performance of meningioma or glioma and narrow intraclass gap, including multidirectional brain region extraction (MDBRE) method and bimodal iterative gamma correction, and combined SE and DenseNet classifier to extract brain regions; the accuracy of triple classification normalized luminance images was improved by 9.7%.

Second, residual SE attention mechanism is to add residual connection to SE branch. Shi et al. [70] combined dense connection with residual SE to propose myopic eye detection network (MDNet), which introduced residual SE correction for dense unit feature mapping channels, reduced the number of parameters, and improved network performance; network detected spherical equivalent mean absolute error up to 1.1150 d (Diopter); feasibility and applicability were verified in fundus images.

Third, multilevel SE attention mechanism is introduced after feature map grouping; personnel reidentification task is critical to capture appearance changes in different viewpoints; localization and alignment of body part are inaccurate, and it is difficult for CNN to capture distinguished information to generate robust representation of pedestrians, for which Yan et al. [71] based on interdependence between different channels and horizontal feature maps, using spatial attention mechanism to learn distinguished local features and construct multilevel SE attention blocks for different subnets, adaptively discover body parts and recalibrate features to solve misalignment problems caused by inaccurate pedestrian detectors.

Fourth, channel spatial attention mechanism is introduced in dense unit, and tandem channel attention and spatial attention constitute convolutional block attention module (CBAM), which can be widely used to improve convolutional neural network representations. Xue et al. [72] proposed dual-attention dense ATP-DenseNet for handwritten font recognition, in which CBAM channel attention aggregates spatial features of average and maximum pools; channel attention map is computed using shared multilayer perceptron; operations along channel axis can effectively highlight information regions; average and maximum pools process features to map channel features and generate a

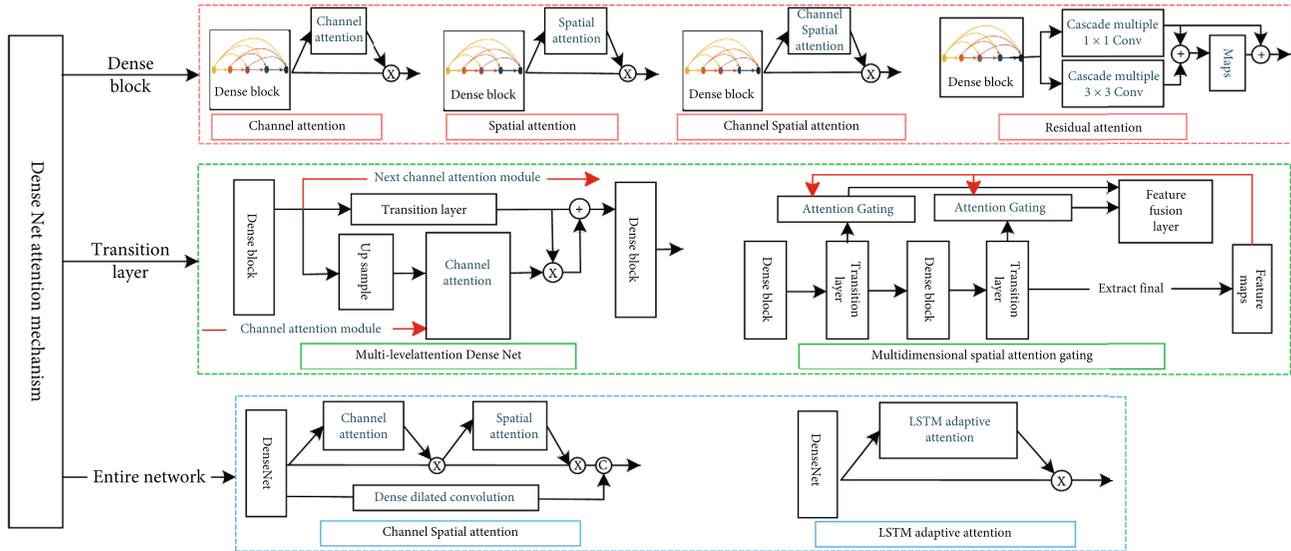


FIGURE 13: Attention mechanism in DenseNet.

two-dimensional spatial attention map in tandem; CBAM is introduced into dense unit to facilitate interlayer information flow and refine features across layers, while target-scale features are given more attention.

Fifth, 3D multiscale attention mechanism is introduced in dense unit. Zhang et al. [73] used 3D densely connected convolutional neural network (CAM-CNN) to extract brain MRI multilevel features for classification of Alzheimer's disease and mild cognitive impairment, densely connected difference at different unit levels; 3D dense unit introduces attention mechanism to generate attention maps and sum transformed MRI hierarchical data into more compact high-level; model has high classification prediction accuracy, and classification performance is at the highest level.

Sixth, convolutional attention mechanism is introduced in dense unit, considers interlayer relationships, and makes full use of interlayer information in convolutional and skip connection. In superresolution tasks, the main difference between high-resolution and low-resolution images lies in edges and textures, which are mainly based on high-frequency feature information to guide image texture recovery, residual superresolution networks have similar convolution and skip processing, resulting in some high-frequency information being missed in network. Bai et al. [74] proposed dense convolutional attention networks (DCAN) to introduce convolutional attention mechanism, which introduces weighted skip connection in convolution, considering interdependence between layers and adaptively calibrating convolutional features to learn richer details and sharper edges by crossing layers; DCAN spends less time to obtain superior performance in both quantity and quality.

3.5.2. Attention Mechanism in DenseNet. DenseNet mainly consists of dense block and transition layer, introducing attention mechanism to enhance feature propagation and reduce the number of network parameters, and introducing attention mechanism in DenseNet is shown in Figure 13.

(1) *Attention Mechanism Introduced in Dense Block.* Attention mechanism introduced in dense block has channel, spatial, channel and spatial, and residual attention mechanisms.

First is attention mechanism in dense block for feature enhancement. Zhou et al. [75] proposed 3D SE-DenseNet for dynamically enhancing hepatocellular carcinoma grading in MR images, introducing SE to enhance key features while suppressing redundant features, feature maps are easier to fuse, and network obtains better performance and accuracy. Squeeze excitation simulates interdependencies between channels but has limited processing of contextual information, Wang et al. [76] used dilated dense unit to extract multiscale features for crowd counting, channel attention mechanism to guide dense block multiscale contextual feature fusion, adjustment of feature channel weights at each fusion stage, recalibrate of fused feature responses, and its convolution layer to capture channel correlation and global average pooling aggregated output feature spatial information; bootstrap channel fusion to adjust feature method outperforms other methods in population density map estimation.

Second, spatial attention mechanism is introduced in dense block. Wei et al. [77] proposed Att-DenseUnet for skin lesion segmentation; spatial attention module was added to encoding and decoding skip connection, using advanced semantic contextual information in encoding stage to focus on lesion region features and suppressing irrelevant artifact region features in dense block, and final model output had high shape accuracy, clear boundary, and high recall segmentation results; Zhang et al. [78] used a multiscale model to automatically classify traffic patterns and speeds, and multiscale spatio-temporal granularity grid data features were extracted in dense block; spatial attention map is obtained after element-by-element dot product of 1×1 convolution and sigmoid functions, which reduces the number of network parameters and enhances interchannel feature

fusion and further enhances information flow in dense block using residual connection; final model outperforms existing models in terms of accuracy, recall, and F1 scores.

Third, channel spatial attention is introduced in dense block, which adaptively adjusts cross-channel and spatial region weight to improve network performance using inter-channel and interfeature relationships. Qin and Gu [79] introduced channel and spatial attention to residual dense block and proposed deep multilevel residual attention network (MRAN) with enhanced image high-frequency features. Attention module aggregates channel information with feature statistics and multilayer perceptron compress channel information and sums it pixel by pixel, emphasizes information-rich channels, computes feature statistics to obtain spatial mapping, and multiplies element-by-element, and spatial attention maps along channel axis improve network perception, focusing more on informative rich features to recover more accurate details, and extensive evaluations show that MRAN is quantitative results and visual performance state-of-the-art methods.

Fourth, residual attention mechanism is introduced in dense block to maintain top-down spatial attention without interrupting bottom-up convolutional feature extraction process, and the residual connection attention mechanism fuses two different convolutional features and enhances feature representation. Bi et al. [80] proposed residual attention densely connected RADC for aerial image scene classification, residual attention mechanism highlights local semantics in output features, and experiments show that RADC with fewer parameters outperforms some existing methods.

(2) Attention Mechanism Introduced in Transition Layer. Attention mechanism introduced in transition layer has multilevel attention mechanism and multidimensional spatial attention mechanism with spatial attention gating.

First, multilevel attention mechanism is introduced in transition layer to fuse adjacent dense block output features and generate channel spatial attention map to enhance transition layer features. In diagnosing gastric cancer in gastric pathology section images, although DenseNet rich semantic information can detect larger sensory field data, it is still difficult to obtain spatial information and identify hidden gastric features in feature maps, Liu et al. [81] proposed multilevel attention dense network (MSA-DenseNet) to detect gastric cancer in 20x magnified section images; transition layer introduced attention providing attention vectors to enhance gastric features, selecting more semantic information about cancer, global average pool, and fully connected layer to provide nonlinearity for channel attention, using adjacent dense block features to enhance relevant semantic information, combining lower-order features with more spatial prediction and higher-order features with more semantic prediction features; final model obtains better detection results than existing manual detection methods.

Second, multidimensional attention with spatial attention gating is introduced between dense block at different scales to adjust multiscale feature fusion and improve network prediction capability. In noninvasive clinical image

lung cancer diagnosis research, which lacks reliable extraction of fine-grained features in different imaging modalities, Qin [82] proposed multidimensional attention-based lung cancer classification architecture for PET and CT images, with 3D DenseNet processing different imaging modalities in parallel, embedding spatial attention gating in dense block, generating corresponding spatial mappings by high-level semantic features and SE, implementing multidimensional attention mechanism, enhancing fine-grained feature extraction, and effectively reducing feature noise; quantitative evaluation metrics and deep learning visualization analysis shows that architecture obtains 0.72 ROC curve area and 0.92 accuracy, which can effectively extract fine-grained features of different imaging modalities.

(3) Attention Mechanism Introduced to Whole Network.

Attention mechanism introduced in whole network has channel spatial and LSTM adaptive attention mechanism. First, skin lesion regions vary in color, size, and shape; some light lesions are highly similar to normal skin, and boundaries of deeper lesion regions are complex, to segment model with sufficiently dense feature resolution, Ren et al. [83] used serial attention network (SANet) to segment skin lesion regions and introduced channel first followed by a spatial attention mechanism to aggregate global, local, and inter-channel information using interdependencies between channel mappings to improve representation of semantic features and using spatial attention to select global contextual information and contextual relationships to make semantic features more compact and consistent; SANet captures interpixel and interchannel feature dependencies and achieves 0.7692 average Jaccard index in ISIC2017 dataset. Second, it is difficult for image caption generation tasks to correctly extract image global features and perform image regions for each word without ignoring partial words, Deng et al. [84] used DenseNet to extract features required for LSTM generation of sentences and introduced LSTM adaptive attention mechanism to improve correspondence problem of forced text and image regions, which was demonstrated on Flickr30k and COCO datasets showed that flexibility of caption generation was improved, with significant improvements in the BLEU and METEOR evaluation criteria.

4. Application of DenseNet in Medical Image Analysis

Reviewing the past 10 years, deep learning has seen a proliferation of improvements in models and algorithms, and availability and computational power of large-scale image data has continued to improve, solving many challenges in computer vision. In 2022, Li et al. [85] provided a comprehensive summary of deep learning research in medical image classification, detection and segmentation, alignment, and retrieval; applications of DenseNet for medical images mainly include image classification, segmentation, detection, alignment, reconstruction, retrieval, generation, enhancement, and fusion, but research and improvement of its technology mainly focus on classification and recognition,

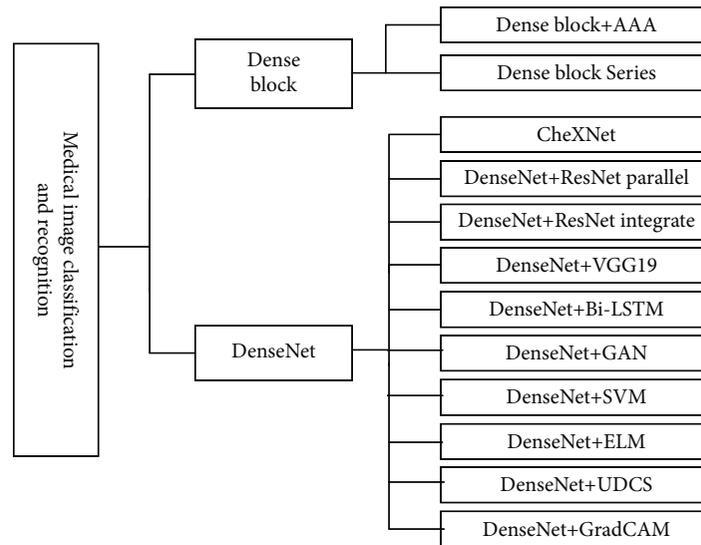


FIGURE 14: Application of DenseNet in medical classification and recognition.

segmentation, and localization detection; medical image classification commonly used application scenarios are disease diagnosis; medical image segmentation projects are mostly applied to segmentation of lesions and organs, such as brain segmentation, lung segmentation, heart segmentation, and liver segmentation. Therefore, this section summarizes applications of DenseNet in field of medical image analysis from these three aspects.

4.1. Application of DenseNet in Medical Pattern Recognition. Image pattern recognition is a popular research area in computer vision, which usually includes two steps: feature extraction and classifier. Medical image classification is further divided into image screening and lesion classification. Image screening uses detection images as model input to predict whether the disease is suffered or graded in severity. Lesion classification assists clinicians in diagnosing human lesions and grading the severity quantitatively, and it is important to identify tumors and nodules in human tissues and organs. This section summarizes applications of medical image pattern recognition from perspective of combining DenseNet with other methods; 12 improved methods are shown in Figure 14.

4.2. Combine Dense Block with Other Methods. First, Guo and Yuan [86] combined dense block and adaptive aggregation attention (AAA) module to classify wireless capsule endoscopy images (WCE); AAA captured global correlation to estimate abnormal regions and obtained 93.17% accuracy in quadruple crossover on CAD-CAP WCE dataset. Second, in pediatric radiology and forensic medicine, bone age assessment is one of clinical diagnostic techniques to evaluate skeletal development in children, Guo et al. [87] tandem six dense blocks to perform comprehensive evaluation of high- and low-quality X-ray images to achieve state-of-the-art performance under MAE metric.

4.3. DenseNet Variants. Rajpurkar et al. [88] proposed 121-layer convolutional neural network CheXNet based on DenseNet, applied to the largest publicly available chest X-ray dataset ChestX-ray14, which contains over 100,000 X-ray images of 14 diseases, and CheXNet exceeded average radiologist $F1$ metrics on a test set annotated by four practicing radiologists performance.

4.4. Combine DenseNet with Others. DenseNet integrates with other networks: first, chest X-ray images, as the most common screening technique for diagnosing chest diseases (atelectasis, cardiomegaly, pneumonia, etc.) suffer from complex background, many potential abnormalities, diverse interactions between abnormal patterns, etc. Traditional techniques focus on low-level features such as appearance, texture, and contrast, although deep networks can extract higher-level features, single network in intermediate layers also loses some unique details, and homogeneous networks learning complementary features from each other may overlap, Chen et al. [89] used DualCheXNet, a dual asymmetric network, to improve performance of multilabel chest disease classification, with parallel asymmetric subnetworks DenseNet and ResNet both using jump connections to create short paths, and enhanced gradient propagation from previous to following layers, residual connection to reuse early features, and dense connection to mine new features; advantages of different unique feature extraction methods complemented each other and were validated on ChestX-ray14 dataset, obtaining average AUC score of 0.823. Second, Srivastva et al. [90] proposed PlexNet integrated with pretrained DenseNet and ResNet for ECG biometrics, feature level fusion approach for tandem features, and fully connected layer based on similarity classification, using both migration learning and integrated learning to take advantage of model effectiveness and robustness demonstrated on PTB and CYBHI datasets. Third, Kedia et al. [91] combined DenseNet121 and VGG19 into CoVNet-19 two-level stacking

model, both models extracted chest X-ray features and minimized noise, SVM achieved three-level classification of COVID-19, pneumonia, and normal, obtaining 98.28% overall accuracy and 97.15% Mathews correlation coefficient and F1 scores of 99% in both binary classifications.

DenseNet combined with other network structures; fourth, Li et al. [92] combined DenseNet and bidirectional recurrent neural network (Bi-LSTM) to predict vivo RNA transcripts from in vitro data; Bi-LSTM learns RNA sequences, secondary structure information and pairwise probability features, which were reshaped into matrices and fed into DenseNet with prediction accuracy and scalability in with better performance. Fifth, Man et al. [93] screened mislabeled plaques in breast histopathology images based on generative adversarial network (GAN) unsupervised abnormality detection; DenseNet extracted multilevel features of plaques for benign and malignant classification, achieving excellent results on coarse-grained high-resolution images especially 40x and 100x images.

DenseNet is combined with other methods; sixth, Devnath et al. [94] used CheXNet to extract multilevel features of chest X-ray images, mixed SVM and CNN aggregation methods, transition layer low-level features, and deep-level features mapped to high-dimensional space after merging dichotomous classification. Seventh, Turkoglu [95] proposed a multicore extreme learning machine (ELM) method to extract deep-level features of chest CT with pretrained DenseNet201, and ELM classifier classified features with different activation methods to predict final presence of neo-coronary pneumonia by majority voting. Eighth, Yuan et al. [96] addressed the problem of imbalance and discrepancy between small and large categories in wireless capsule endoscopy images and proposed DenseNet-UDCS polyp recognition model by calculating suitable features, unbalance discriminant loss (UD) function to improve model sensitivity to minority groups and category-sensitive loss (CS) constraint to learn features in hierarchically between different categories, so that intraclass differences minimized and interclass distance within batches maintained, with more advantages in model recognition accuracy, computational speed, and stability. Ninth, Ellis et al. [97] mixed weakly supervised, extrasupervised, and annotated classification of two-view chest radiographs; DenseNet-121 extracted features to generate 1D global feature vector and estimated scores; fully connected layer gradients were convolved with feature sets to estimate gradient class activation maps (Grad-CAM); attention mining branch learned GradCAM threshold operations using DenseNet-121, which masked important classification regions in input, increasing accuracy and sensitivity to negative prediction values, despite a decrease in false positives and decreases in overall accuracy.

4.5. Application of DenseNet in Medical Image Segmentation. Medical image segmentation is to identify contour of target region or internal voxel collections, which is a key task for clinical surgical images to guide tumor radiotherapy, medical image organ and lesion segmentation allows quantitative analysis of volume and shape-related clinical parameters. Effective automatic organ segmentation is very challenging,

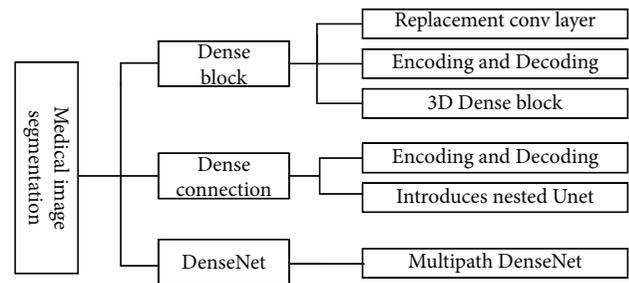


FIGURE 15: Application of DenseNet in medical image segmentation.

medical images have high complexity, lack of simple linear features, insignificant brightness differences between organs, connected to surrounding organs, and blurred boundaries, as well as interference factors such as image artifacts and noise, and segmentation results accuracy is also affected by partial volume effects, gray scale inhomogeneity, artifacts and gray scale proximity between different soft tissues. Traditional segmentation algorithms are difficult to process images accurately, and application of DenseNet makes medical image segmentation effect improved, summarizing the seven models shown in Figure 15.

4.5.1. Dense Block in Medical Image Segmentation Application Research. First, dense block replaces convolutional layer. Ding et al. [98] used dense block instead of codec convolutional layer to extract low-level visual features of multimodal brain tumors, and fused high-level semantic features to generate high-resolution features with fewer network parameters and fast segmentation, achieving good segmentation results. AIGhamdi et al. [99] introduced dense block in U-net for breast artery calcification detection; DU-net has dense connectivity, which helps to improve computational reusability and gradient mobility and improves accuracy and training difficulty. Second, codec introduces dense block. Vila et al. [100] performed semantic segmentation of carotid ultrasound plaques based on DenseNet, dense connectivity captured multiscale contextual information, and correlation coefficient of carotid intima-media thickness reached 0.81 in the experiment; robustness and generalization ability were validated on two datasets. Thirdly, 3D dense block was applied to segmentation networks. Bui et al. [101] proposed 3D-SkipDenseSeg, an accurate segmentation method for infant brain MRI, in which jump connections directly combine fine and coarse feature information from different levels of dense block to alleviate low accuracy problem caused by low-intensity contrast between MRI tissues, final model obtained the best dice similarity coefficient of 90.37, and segmentation accuracy and parameter efficiency are high.

(1) Dense Connection in Medical Image Segmentation Application Research. First, dense connection is introduced in codecs. Shi et al. [102] proposed dual dense context-aware network (DDCNet) to automatically segment hippocampus; dense connection is introduced between codecs to fuse high level and low level multiresolution features and contextual

information, which can enhance interlayer feature propagation and improve information flow between codecs. Second, nested U-net introduces dense connection; Li et al. [103] proposed ANU-Net combining deeply supervised codec structure and nested segmentation network, designed full resolution nested U-net between codecs, introduced dense connection between nested convolutional layers of same resolution to obtain full-resolution features at different semantic levels, and experimentally demonstrated that ANU-Net can suppress irrelevant background regions while increase target region weight.

(2) *DenseNet in Medical Image Segmentation Application Research*. Multipath DenseNet extracts features. Kitrungrotsakul et al. [104] proposed multipath DenseNet for hepatic vascular segmentation, binary classification of blocks in three planes, sagittal, coronal and transverse; IRCAD dataset experiments show that the model is currently the latest method accuracy.

4.6. Application of DenseNet in Medical Object Detection. Localization of anatomical structures and lesion regions in human organs is an important preprocessing step in clinical treatment planning and intervention, and localization of specific biomarkers or anatomical structures in medical images is directly related to treatment outcome. The key target of interest or lesion detection is to classify each pixel; DenseNet is applied to improve the effectiveness of medical image localization detection, with the following three main models.

First, Ghatwary et al. [105] detected esophageal abnormalities in endoscopic images based on faster R-CNN- and DenseNet-enhanced feature propagation between layers, combined with Gabor manual feature enhancement to detect texture details in stage, obtaining 92.1% and 91% accuracy on MICCAI2015 and Kvasir datasets. Second, Liu et al. [106] proposed fully convolutional dense pixel classifier (FC-DenseNet) depicting ribs and clavicles, aggregating contextual information with jump connections, dense blocks instead of upsampling paths in convolutional layer, and dense connections to enhance feature propagation and reuse, and FC-DenseNet automatically depicts chest radiographs from multiple public databases to generate accurate binary edge maps, verifying the effectiveness of the model for bone suppression. Third, Xiao et al. [107] designed Siam network based on DenseNet to track vessel wall displacement from ultrasound RF signals and target blocks of labeled data as convolutional kernels to calculate blocks in search region with similarity score maps, achieving first single-target displacement tracking in ultrasound pulse wave imaging, accurately predicting local mechanical properties of arteries, and tracking better than traditional single-target block matching better.

5. Summary and Outlook

In summary, DenseNet, which can learn deeper and more discriminative features from images, has been applied to several research areas of medical image analysis and has made

breakthroughs; although DenseNet has been effectively validated in publicly available and relevant medical image databases [41], it still faces many challenges. First, the huge amount of medical data outside images [4] contains rich information, but it is difficult to be fully utilized, such as additional labels indicating lesion conditions, clinical diagnosis reports reflecting physicians' conclusions and descriptions, which can help deep learning models to improve diagnostic performance, and recurrent convolutional neural networks (RNN) can be utilized to incorporate medical other data into DenseNet.

Second, deep learning models rely on a large number of manually labeled datasets [108, 109], which usually require significant effort and time to collect, clean, and debug data, and changes and evolution of actual task requirements also lead to retagging of the datasets, data dependence of models and cost of dataset labeling are issues that need to be urgently addressed by researchers. With the limited labeled trainable data [110, 111] currently available for training, semisupervised and weakly supervised methods can be used to learn unlabeled, weak, and small portions of labeled data.

Finally, there are still many challenges to training a model with higher classification or prediction accuracy and better generalization ability. There are limitations in using multiple branches to achieving DenseNet broaden; DenseNet lightweight is not only studied from the direction of architecture design and model compression but also based on data considerations, such as dimensionality reduction of intermediate features; dense connections lead to redundancy of feature extraction in different layers; dense block has excellent performance, but transition layer structure of dimensionality reduction is too simple, and it is difficult for single sensory field to capture multilevel features of dense block; the existing attention mechanism is basically in form of "weighting," and appropriate attention mechanism should be designed according to data characteristics, such as residual shrinkage network [112]; "soft thresholding" is used to cope with strong noisy data.

Massive medical domain data presents challenges and opportunities for researchers in medical image analysis, by appropriate integration methods, different types of domain knowledge are used to assist deep learning to better accomplish corresponding tasks, and natural image pretraining weights are also introduced to improve network performance in medical domain. The massive data in medical image analysis will provide inexhaustible power as network architecture and algorithms improve; dense block and dense connection in DenseNet can be optimized for these new models, and DenseNet-base derivative models may also yield significant improvements in results for specific tasks.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (grant no. 62062003), Key Research and Development Project of Ningxia (grant no. 2020BEB04022), Natural Science Foundation of Ningxia (grant no. 2022AAC03149), and North Minzu University Research Project of Talent Introduction (grant no. 2020KYQD08).

References

- [1] I. S. Krizhevsky and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceeding of Neural Information Processing Systems*, pp. 1097–1105, Curran Associates, Lake Tahoe, USA, 2012.
- [2] G. Huang, Z. Liu, V. Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," 2016, <https://arxiv.org/abs/1608.06993>.
- [3] G. Litjens, T. Kooi, B. E. Bejnordi et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [4] X. Xie, J. Niu, X. Liu, Z. Chen, S. Tang, and S. Yu, "A survey on incorporating domain knowledge into deep learning for medical image analysis," *Medical Image Analysis*, vol. 69, 2021.
- [5] J. He, T. Cao, F. Xu et al., "Artificial intelligence-based screening for diabetic retinopathy at community hospital," *Eye*, vol. 34, no. 3, pp. 572–576, 2020.
- [6] D. Wang, C. Huang, S. Bao et al., "Study on the prognosis predictive model of COVID-19 patients based on CT radiomics," *Scientific Reports*, vol. 11, no. 1, 2021.
- [7] R. Zhang, L. Zhao, T. Lou et al., "Automatic segmentation of acute ischemic stroke from DWI using 3-D fully convolutional DenseNets," *IEEE Transactions on Medical Imaging*, vol. 37, no. 9, pp. 2149–2160, 2018.
- [8] F. Li, M. Liu, and Alzheimer's Disease Neuroimaging Initiative, "Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks," *Computerized Medical Imaging and Graphics*, vol. 70, pp. 101–110, 2018.
- [9] S. Jégou, M. Drozdal, D. Vazquez, A. Romero, and Y. Bengio, "The one hundred layers tiramisu: fully convolutional densenets for semantic segmentation," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, vol. 17, pp. 1175–1183, Honolulu, HI, USA, 2017.
- [10] L. Ke, Y. Deng, W. Xia et al., "Development of a self-constrained 3D DenseNet model in automatic detection and segmentation of nasopharyngeal carcinoma using magnetic resonance images," *Oral Oncology*, vol. 110, 2020.
- [11] Z. Liu, Y. Zhang, W. Li, S. Li, Z. Zou, and B. Chen, "Multislice left ventricular ejection fraction prediction from cardiac MRIs without segmentation using shared SptDenNet," *Computerized Medical Imaging and Graphics*, vol. 86, 2020.
- [12] D. M. H. Kishan and V. Janardhan, "Intracranial hemorrhage detection," *Materials Today: Proceedings*, vol. 43, no. 6, pp. 3706–3714, 2021.
- [13] L. Li, Y. Yang, H. Liang, and B. Wu, "Transfer learning for establishment of recognition of COVID-19 on CT imaging using small-sized training datasets," *Knowledge-Based Systems*, vol. 218, article 106849, 2021.
- [14] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, 2016.
- [16] Y. Chen, J. Li, H. Xiao, X. Jin, S. Yan, and J. Feng, "Dual path networks," 2017, <https://arxiv.org/abs/1707.01629>.
- [17] Y. Yang, Z. Zhong, T. Shen, and Z. Lin, "Convolutional Neural Networks with Alternately Updated Clique," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2413–2422, Salt Lake City, USA, 2018.
- [18] L. J. Ba and R. Caruana, "Do Deep Nets Really Need to Be Deep?," 2013, <https://arxiv.org/abs/1312.6184>.
- [19] B. Hanin and D. Rolnick, "Complexity of Linear Regions in Deep Networks," in *Proceedings of the 11th Asian Conference on Machine Learning (ACML)*, pp. 2596–2604, Nagoya, Japan, 2019.
- [20] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, Las Vegas, NV, USA, 2016.
- [21] H. Li, S. Zhuang, D. Li, J. Zhao, and Y. Ma, "Benign and malignant classification of mammogram images based on deep learning," *Biomedical Signal Processing and Control*, vol. 51, pp. 347–354, 2019.
- [22] T. W. Arega and S. Bricq, "Automatic myocardial scar segmentation from multi-sequence cardiac MRI using fully convolutional densenet with inception and squeeze-excitation module," in *Proceedings of the 23rd International Conference on Medical Image Computing and Computer Assisted Intervention*, Lima, Peru, 2020.
- [23] B. Lodhi and J. Kang, "Multipath-DenseNet: a supervised ensemble architecture of densely connected convolutional networks," *Information Sciences*, vol. 482, pp. 63–72, 2019.
- [24] B. Chen, T. Zhao, J. Liu, and L. Lin, "Multipath feature recalibration DenseNet for image classification," *International Journal of Machine Learning and Cybernetics*, vol. 12, no. 3, pp. 651–660, 2021.
- [25] Y. Liu, P. Hao, P. Zhang, X. Xu, J. Wu, and W. Chen, "Dense convolutional binary-tree networks for lung nodule classification," *IEEE Access*, vol. 6, pp. 49080–49088, 2018.
- [26] M. T. Nguyen, P. Siritanawan, and K. Kotani, "Saliency detection in human crowd images of different density levels using attention mechanism," *Signal Processing: Image Communication*, vol. 88, 2020.
- [27] G. Huang, S. Liu, L. Maaten, and K. Q. Weinberger, "CondenseNet: An Efficient DenseNet Using Learned Group Convolutions," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2752–2761, Salt Lake City, USA, 2018.
- [28] G. Li, M. Zhang, J. Li, F. Lv, and G. Tong, "Efficient densely connected convolutional neural networks," *Pattern Recognition*, vol. 109, article 107610, 2021.
- [29] A. G. Howard, M. Zhu, B. Chen et al., "Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications," 2017, <https://arxiv.org/abs/1704.04861>.

- [30] J. Chen, W. Wang, D. Zhang, A. Zeb, and Y. A. Nanekharan, "Attention embedded lightweight network for maize disease recognition," *Plant Pathology*, vol. 70, no. 3, pp. 630–642, 2021.
- [31] X. Yu, H. He, H. Ma, X. Qi, J. Lu, and Y. Zhao, "S-DenseNet: a DenseNet compression model based on convolution grouping strategy using skyline method," *IEEE Access*, vol. 7, pp. 183604–183613, 2019.
- [32] Z. Huang, X. Zhu, M. Ding, and X. Zhang, "Medical image classification using a light-weighted hybrid neural network based on PCANet and DenseNet," *IEEE Access*, vol. 8, pp. 24697–24712, 2020.
- [33] Y. Lee, J. W. Hwang, S. Lee, Y. Bae, and J. Park, "An Energy and GPU-Computation Efficient Backbone Network for Real-Time Object Detection," in *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 752–760, Long Beach, USA, 2019.
- [34] T. Li, W. Jiao, L. Wang, and G. Zhong, "Automatic DenseNet sparsification," *IEEE Access*, vol. 8, pp. 62561–62571, 2020.
- [35] X. Zhu, Z. Li, X. Zhang, H. Li, Z. Xue, and L. Wang, "Generative adversarial image super-resolution through deep dense skip connections," *Computer Graphics Forum*, vol. 37, no. 7, pp. 289–300, 2018.
- [36] Y. Wei and X. Liu, "Dangerous goods detection based on transfer learning in X-ray images," *Neural Computing and Applications*, vol. 32, no. 12, pp. 8711–8724, 2020.
- [37] X. Wei, D. Wei, D. Suo, L. Jia, and Y. Li, "Multi-target defect identification for railway track line based on image processing and improved YOLOv3 model," *IEEE Access*, vol. 8, pp. 61973–61988, 2020.
- [38] B. Lei, Z. Xia, F. Jiang et al., "Skin lesion segmentation via generative adversarial networks with dual discriminators," *Medical Image Analysis*, vol. 64, article 101716, 2020.
- [39] M. K. Hasan, L. Dahal, P. N. Samarakoon, F. I. Tushar, and R. Martí, "DSNet: automatic dermoscopic skin lesion segmentation," *Computers in Biology and Medicine*, vol. 120, 2020.
- [40] H. T. Mustafa, M. Zareapoor, and J. Yang, "MLDNet: multi-level dense network for multi-focus image fusion," *Signal Processing: Image Communication*, vol. 85, 2020.
- [41] X. Li, L. Shen, X. Xie et al., "Multi-resolution convolutional networks for chest X-ray radiograph based lung nodule detection," *Artificial Intelligence in Medicine*, vol. 103, 2020.
- [42] Z. Feng, J. Lai, X. Xie, and J. Zhu, "Image super-resolution via a densely connected recursive network," *Neurocomputing*, vol. 316, pp. 270–276, 2018.
- [43] L. Yang, Y. Qin, and X. Zhang, "Lightweight densely connected residual network for human pose estimation," *Journal of Real-Time Image Processing*, vol. 18, no. 1, pp. 825–837, 2021.
- [44] M. Wang, X. Yang, M. Anisetti, R. Zhang, M. K. Albertini, and K. Liu, "Image super-resolution via enhanced multi-scale residual network," *Journal of Parallel and Distributed Computing*, vol. 152, pp. 57–66, 2021.
- [45] Y. Wu, Y. Qin, Z. Wang, X. Ma, and Z. Cao, "Densely pyramidal residual network for UAV-based railway images dehazing," *Neurocomputing*, vol. 371, pp. 124–136, 2020.
- [46] Y. Long, H. Jia, Y. Zhong, Y. Jiang, and Y. Jia, "RXDNFuse: a aggregated residual dense network for infrared and visible image fusion," *Information Fusion*, vol. 69, pp. 128–141, 2021.
- [47] B. Zhang, S. Liu, S. Zhou et al., "Cerebrovascular segmentation from TOF-MRA using model- and data-driven method via sparse labels," *Neurocomputing*, vol. 380, pp. 162–179, 2020.
- [48] X. Wang, Y. Gu, X. Gao, and Z. Hui, "Dual residual attention module network for single image super resolution," *Neurocomputing*, vol. 364, pp. 269–279, 2019.
- [49] W. Zeng, J. Peng, S. Wang, and Q. Liu, "A comparative study of CNN-based super-resolution methods in MRI reconstruction and its beyond," *Signal Processing: Image Communication*, vol. 81, 2020.
- [50] Z. Zhang, C. Wu, S. Coleman, and D. Kerr, "DENSE-INception U-net for medical image segmentation," *Computer Methods and Programs in Biomedicine*, vol. 192, 2020.
- [51] S. Qamar, P. Ahmad, and L. Shen, "Dense encoder-decoder-based architecture for skin lesion segmentation," *Cognitive Computation*, vol. 13, no. 2, pp. 583–594, 2021.
- [52] Z. Zhang, Z. Tang, Y. Wang et al., "Dense residual network: enhancing global dense feature flow for character recognition," *Neural Networks*, vol. 139, pp. 77–85, 2021.
- [53] A. Adegun and S. Viriri, "FCN-based DenseNet framework for automated detection and classification of skin lesions in dermoscopy images," *IEEE Access*, vol. 8, pp. 150377–150396, 2020.
- [54] Y. Hong, B. Wei, Z. Han, X. Li, Y. Zheng, and S. Li, "MMCL-Net: spinal disease diagnosis in global mode using progressive multi-task joint learning," *Neurocomputing*, vol. 399, pp. 307–316, 2020.
- [55] Z. Huang, Y. Zhao, X. Li et al., "Application of innovative image processing methods and AdaBound-SE-DenseNet to optimize the diagnosis performance of meningiomas and gliomas," *Biomedical Signal Processing and Control*, vol. 59, article 101926, 2020.
- [56] S. Wang, Z. Liu, Y. Rong et al., "Deep learning provides a new computed tomography-based prognostic biomarker for recurrence prediction in high-grade serous ovarian cancer," *Radiotherapy and Oncology*, vol. 132, pp. 171–177, 2019.
- [57] A. Pemasiri, K. Nguyen, S. Sridharan, and C. Fookes, "Multi-modal semantic image segmentation," *Computer Vision and Image Understanding*, vol. 202, 2021.
- [58] T. J. Jun, J. Kweon, Y. Kim, and D. Kim, "T-Net: nested encoder-decoder architecture for the main vessel segmentation in coronary angiography," *Neural Networks*, vol. 128, pp. 216–233, 2020.
- [59] X. Yang, Y. Zhang, Y. Guo, and D. Zhou, "An image super-resolution deep learning network based on multi-level feature extraction module," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 7063–7075, 2021.
- [60] B. N. Mostefa, A. Moharmed, S. Rachida, and K. Rostom, "Fully automatic brain tumor segmentation with deep learning-based selective attention using overlapping patches and multi-class weighted cross-entropy," *Medical Image Analysis*, vol. 63, article 101692, 2020.
- [61] Y. Zhou, Y. Zhang, X. Xie, and S. Kung, "Image super-resolution based on dense convolutional auto-encoder blocks," *Neurocomputing*, vol. 423, pp. 98–109, 2021.
- [62] A. Signoroni, M. Savardi, S. Benini et al., "BS-Net: learning COVID-19 pneumonia severity on a large chest X-ray dataset," *Medical Image Analysis*, vol. 71, article 102046, 2021.

- [63] K. Tseng, R. Zhang, C. Chen, and M. M. Hassan, "DNetUnet: a semi-supervised CNN of medical image segmentation for super-computing AI service," *The Journal of Supercomputing*, vol. 77, no. 4, pp. 3594–3615, 2021.
- [64] J. Liu, M. Kocak, M. Supanich, and J. Deng, "Motion artifacts reduction in brain MRI by means of a deep residual network with densely connected multi-resolution blocks (DRN-DCMB)," *Magnetic Resonance Imaging*, vol. 71, pp. 69–79, 2020.
- [65] M. Wang, H. Cai, J. Zhou, and M. Zhou, "Interlayer and intralayer scale aggregation for scale-invariant crowd counting," *Neurocomputing*, vol. 441, pp. 128–137, 2021.
- [66] K. Sun, X. Wen, L. Yuan, and H. Xu, "Dense capsule networks with fewer parameters," *Soft Computing*, vol. 25, no. 10, pp. 6927–6945, 2021.
- [67] D. Wang, G. Hu, and C. Lyu, "Multi-path connected network for medical image segmentation," *Journal of Visual Communication and Image Representation*, vol. 71, article 102852, 2020.
- [68] A. Vaswani, N. Shazeer, N. Parmar et al., "Attention is all you need," 2017, <https://arxiv.org/abs/1706.03762>.
- [69] J. Bastings and K. Filippova, "The Elephant in the Interpretability Room: Why Use Attention as Explanation when we Have Saliency Methods?," 2020, <https://arxiv.org/abs/2010.05607>.
- [70] Z. Shi, T. Wang, Z. Huang, F. Xie, and G. Song, "A method for the automatic detection of myopia in Optos fundus images based on deep learning," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 37, no. 6, article e3460, 2021.
- [71] Y. Yan, B. Ni, J. Liu, and X. Yang, "Multi-level attention model for person re-identification," *Pattern Recognition Letters*, vol. 127, pp. 156–164, 2019.
- [72] S. Xue, D. G. Liu, and Y. Ma, "ATP-DenseNet: a hybrid deep learning-based gender identification of handwriting," *Neural Computing and Applications*, vol. 33, no. 10, pp. 4611–4622, 2021.
- [73] J. Zhang, B. Zheng, A. Gao, X. Feng, D. Liang, and X. Long, "A 3D densely connected convolution neural network with connection-wise attention mechanism for Alzheimer's disease classification," *Magnetic Resonance Imaging*, vol. 78, pp. 119–126, 2021.
- [74] W. Bai, Y. Lu, Y. Huang, L. Zha, and J. Yang, "Densely convolutional attention network for image super-resolution," *Neurocomputing*, vol. 368, pp. 25–33, 2019.
- [75] Q. Zhou, Z. Zhou, C. Chen et al., "Grading of hepatocellular carcinoma using 3D SE-DenseNet in dynamic enhanced MR images," *Computers in Biology and Medicine*, vol. 107, pp. 47–57, 2019.
- [76] L. Wang, Y. Li, S. Peng, X. Tang, and B. Yin, "Multi-level feature fusion network for crowd counting," *IET Computer Vision*, vol. 15, no. 1, pp. 60–72, 2021.
- [77] Z. Wei, H. Song, L. Chen, Q. Li, and G. Han, "Attention-based DenseUnet network with adversarial training for skin lesion segmentation," *IEEE Access*, vol. 7, pp. 136616–136629, 2019.
- [78] R. Zhang, P. Xie, C. Wang, G. Liu, and S. Wan, "Classifying transportation mode and speed from trajectory data via deep multi-scale learning," *Computer Networks*, vol. 162, article 106861, 2019.
- [79] D. Qin and X. Gu, "Single-image super-resolution with multilevel residual attention network," *Neural Computing and Applications*, vol. 32, no. 19, pp. 15615–15628, 2020.
- [80] Q. Bi, K. Qin, H. Zhang, Z. Li, and K. Xu, "RADNet: a residual attention based convolution network for aerial scene classification," *Neurocomputing*, vol. 377, pp. 345–359, 2020.
- [81] B. Liu, Y. Zhao, B. Yang, S. Zhao, R. Gu, and M. Gahegan, "A gastric cancer recognition algorithm on gastric pathological sections based on multistage attention-DenseNet," *Concurrency and Computation Practice and Experience*, vol. 33, no. 10, article e6188, 2021.
- [82] R. Qin, Z. Wang, L. Jiang et al., "Fine-grained lung cancer classification from PET and CT images based on multidimensional attention mechanism," *Complexity*, vol. 2020, Article ID 6153657, 12 pages, 2020.
- [83] Y. Ren, L. Yu, S. Tian, J. Cheng, Z. Guo, and Y. Zhang, "Serial attention network for skin lesion segmentation," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, pp. 799–810, 2021.
- [84] Z. Deng, Z. Jiang, R. Lan, W. Huang, and X. Luo, "Image captioning using DenseNet network and adaptive attention," *Signal Processing: Image Communication*, vol. 85, 2020.
- [85] X. Li, C. Li, M. M. Rahaman et al., "A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches," *Artificial Intelligence Review*, vol. 29, pp. 609–639, 2022.
- [86] X. Guo and Y. Yuan, "Semi-supervised WCE image classification with adaptive aggregated attention," *Medical Image Analysis*, vol. 64, 2020.
- [87] J. Guo, J. Zhu, H. Du, and B. Qiu, "A bone age assessment system for real-world X-ray images based on convolutional neural networks," *Computers & Electrical Engineering*, vol. 81, 2020.
- [88] P. Rajpurkar, J. Irvin, K. Zhu et al., "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," 2017, <https://arxiv.org/abs/1711.05225>.
- [89] B. Chen, J. Li, X. Guo, and G. Lu, "DualCheXNet: dual asymmetric feature learning for thoracic disease classification in chest X-rays," *Biomedical Signal Processing and Control*, vol. 53, article 101554, 2019.
- [90] R. Srivastava, A. Singh, and Y. N. Singh, "PlexNet: a fast and robust ECG biometric system for human recognition," *Information Sciences*, vol. 558, pp. 208–228, 2021.
- [91] P. Kedia, B. Anjum, and R. Katarya, "CoVNet-19: a deep learning model for the detection and analysis of COVID-19 patients," *Applied Soft Computing*, vol. 104, article 107184, 2021.
- [92] Z. Li, J. Zhu, X. Xu, and Y. Yao, "RDense: a protein-RNA binding prediction model based on bidirectional recurrent neural network and densely connected convolutional networks," *IEEE Access*, vol. 8, pp. 14588–14605, 2020.
- [93] R. Man, P. Yang, and B. Xu, "Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks," *IEEE Access*, vol. 8, pp. 155362–155377, 2020.
- [94] L. Devnath, S. Luo, P. Summons, and D. Wang, "Automated detection of pneumoconiosis with multilevel deep features learned from chest X-Ray radiographs," *Computers in Biology and Medicine*, vol. 129, 2021.

- [95] M. Turkoglu, "COVID-19 detection system using chest CT images and multiple kernels-extreme learning machine based on deep neural network," *IRBM*, vol. 42, no. 4, pp. 207–214, 2021.
- [96] Y. Yuan, W. Qin, B. Ibragimov et al., "Densely connected neural network with unbalanced discriminant and category sensitive constraints for polyp recognition," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 2, pp. 574–583, 2020.
- [97] R. Ellis, E. Ellestad, B. Elicker, M. D. Hope, and D. Tosun, "Impact of hybrid supervision approaches on the performance of artificial intelligence for the classification of chest radiographs," *Computers in Biology and Medicine*, vol. 120, article 103699, 2020.
- [98] Y. Ding, L. Gong, M. Zhang, C. Li, and Z. Qin, "A multi-path adaptive fusion network for multimodal brain tumor segmentation," *Neurocomputing*, vol. 412, pp. 19–30, 2020.
- [99] M. AlGhamdi, M. Abdel-Mottaleb, and F. Collado-Mesa, "DU-Net: convolutional network for the detection of arterial calcifications in mammograms," *IEEE Transactions on Medical Imaging*, vol. 39, no. 10, pp. 3240–3249, 2020.
- [100] M. D. M. Vila, B. Remeseiro, M. Grau et al., "Semantic segmentation with DenseNets for carotid artery ultrasound plaque segmentation and CIMT estimation," *Artificial Intelligence in Medicine*, vol. 103, 2020.
- [101] T. D. Bui, J. Shin, and T. Moon, "Skip-connected 3D DenseNet for volumetric infant brain MRI segmentation," *Biomedical Signal Processing and Control*, vol. 54, article 101613, 2019.
- [102] J. Shi, R. Zhang, L. Guo et al., "Dual dense context-aware network for hippocampal segmentation," *Biomedical Signal Processing and Control*, vol. 61, article 102038, 2020.
- [103] C. Li, Y. Tan, W. Chen et al., "ANU-Net: attention-based nested U-Net to exploit full resolution features for medical image segmentation," *Computers & Graphics*, vol. 90, pp. 11–20, 2020.
- [104] T. Kitrungratsakul, X. Han, Y. Iwamoto et al., "VesselNet: a deep convolutional neural network with multi pathways for robust hepatic vessel segmentation," *Computerized Medical Imaging and Graphics*, vol. 75, pp. 74–83, 2019.
- [105] N. Ghatwary, X. Ye, and M. Zolgharni, "Esophageal abnormality detection using DenseNet based faster R-CNN with Gabor features," *IEEE Access*, vol. 7, pp. 84374–84385, 2019.
- [106] Y. Liu, X. Zhang, G. Cai et al., "Automatic delineation of ribs and clavicles in chest radiographs using fully convolutional DenseNets," *Computer Methods and Programs in Biomedicine*, vol. 180, article 105014, 2019.
- [107] C. Xiao, Z. Li, J. Lu et al., "A new deep learning method for displacement tracking from ultrasound RF signals of vascular walls," *Computerized Medical Imaging and Graphics*, vol. 87, article 101819, 2021.
- [108] Z. Tao, L. Huiling, W. Wenwen, and Y. Xia, "GA-SVM based feature selection and parameter optimization in hospitalization expense modeling," *Applied Soft Computing*, vol. 75, pp. 323–332, 2019.
- [109] M. M. Rahaman, C. Li, Y. Yao, and X. Zhao, "Identification of COVID-19 samples from chest X-ray images using deep learning: a comparison of transfer learning approaches," *Journal of X-Ray Science and Technology*, vol. 28, no. 5, pp. 821–839, 2020.
- [110] Y. Li, C. Li, X. Li et al., "A comprehensive review of Markov random field and conditional random field approaches in pathology image analysis," *Archives of Computational Methods in Engineering*, vol. 29, no. 1, pp. 609–639, 2022.
- [111] T. Zhou, H. Lu, Z. Yang, S. Qiu, B. Huo, and Y. Dong, "The ensemble deep learning model for novel COVID-19 on CT images," *Applied Soft Computing*, vol. 98, article 106885, 2021.
- [112] M. Zhao, S. Zhong, X. Fu, B. Tang, and M. Pecht, "Deep residual shrinkage networks for fault diagnosis," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4681–4690, 2020.