Computers in Biology and Medicine 153 (2023) 106496

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Multi-task deep learning for medical image computing and analysis: A review

Yan Zhao [a], Xiuying Wang [b,**], Tongtong Che [a], Guoqing Bao [b], Shuyu Li [c,*]

[a] Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing, 100083, China
[b] School of Computer Science, The University of Sydney, Sydney, NSW, 2008, Australia
[c] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, 100875, China

ARTICLE INFO

Keywords:
Deep learning
Multi-task learning
Medical image application
Medical image analysis
Survey

ABSTRACT

The renaissance of deep learning has provided promising solutions to various tasks. While conventional deep learning models are constructed for a single specific task, multi-task deep learning (MTDL) that is capable to simultaneously accomplish at least two tasks has attracted research attention. MTDL is a joint learning paradigm that harnesses the inherent correlation of multiple related tasks to achieve reciprocal benefits in improving performance, enhancing generalizability, and reducing the overall computational cost. This review focuses on the advanced applications of MTDL for medical image computing and analysis. We first summarize four popular MTDL network architectures (i.e., cascaded, parallel, interacted, and hybrid). Then, we review the representative MTDL-based networks for eight application areas, including the brain, eye, chest, cardiac, abdomen, musculoskeletal, pathology, and other human body regions. While MTDL-based medical image processing has been flourishing and demonstrating outstanding performance in many tasks, in the meanwhile, there are performance gaps in some tasks, and accordingly we perceive the open challenges and the perspective trends. For instance, in the 2018 Ischemic Stroke Lesion Segmentation challenge, the reported top dice score of 0.51 and top recall of 0.55 achieved by the cascaded MTDL model indicate further research efforts in high demand to escalate the performance of current models.

1. Introduction

Medical imaging [1] plays an increasingly crucial role in modern medicine. It has been routinely and widely used for versatile clinical practices, ranging from disease detection (e.g., detecting the location of stroke lesion) and treatment planning to image-guided surgery and disease prognosis monitoring. The medical imaging modalities, including X-ray, ultrasound, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET), and more specific modalities for particular organs such as mammography, colonoscopy, retinal fundus photography, reflect changes due to various medical conditions from the structural, functional, or metabolic levels. These imaging data account for about 90% of overall healthcare data [2]. Processing and interpreting these large-scale, complex, and diverse medical images efficiently and precisely and further effectively mining and identifying clinically meaningful patterns are far beyond human capability and capacity. Hence, there has been a consistent and pro-longed demand for automated and intelligent computational solutions.

Artificial intelligence has been proven to alleviate the challenges with its outperforming capacities in computing and analyzing the overwhelming amounts of images [3]. Conventional machine learning requires human engineering and domain expertise to structure data and design feature extractors. In contrast, deep learning (DL) [4] enables end-to-end learning of very complex functions or intricate representations from raw high-dimensional data. Benefitting from the increased computing power and open available large-labeled datasets, DL-related algorithms have gained astonishing progress in the last two decades and closely approached or even surpassed human-level performance for general computer vision tasks [5,6]. DL has also led to a revolution in the medical image community [7,8]. Many DL models have been exploited to complete different medical image computing and analysis tasks, enhancing computing speed and performance accuracy significantly.

Most medical image computing and analysis research has focused on

* Corresponding author. State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, 100875, China.
** Corresponding author.
E-mail addresses: xiu.wang@sydney.edu.au (X. Wang), shuyuli@bnu.edu.cn (S. Li).

https://doi.org/10.1016/j.compbiomed.2022.106496
Received 21 September 2022; Received in revised form 6 December 2022; Accepted 27 December 2022
Available online 28 December 2022

separately and independently solving various tasks, i.e., one model for one task. However, there exist inherent or complementary relationships among these tasks. For example, the lung cancer diagnosis based on CT relies on nodule-level segmentation and classification [9]; there exists a commensal correlation between bi-ventricle segmentation and direct area estimation [10]. To thoroughly study the disease from multiple views, some available medical image datasets have provided different types of annotations. For example, the brain tumor segmentation (BraTS) 2021 challenge[1] has provided segmentation labels of the histologically distinct brain tumor sub-regions, as well as the methylation status; The Alzheimer's disease neuroimaging initiative (ADNI) dataset[2] has identified the disease status and measured several cognitive scores. Recently, to take full advantage of inherent relations among tasks and different kinds of medical image annotations, increasing efforts have attempted to accomplish multiple tasks by jointly training a single model under the multi-task deep learning (MTDL) paradigm, which can also be referred to as multitask deep learning or deep multi-task learning.

Different from the conventional deep learning that completes each task in isolation, MTDL aims to tackle multiple related tasks by simultaneously optimizing the loss functions of multiple tasks. Mathematically, given an input $\boldsymbol{X}$, MTDL aims to help improve the joint learning of multiple tasks $\{T_i\}_{i=1}^m$ with a deep learning-based model $\boldsymbol{F}$ and output the corresponding targets $\{\boldsymbol{Y_i}\}_{i=1}^m$, which can be formulated as $\{\boldsymbol{Y_i}\}_{i=1}^m = \boldsymbol{F}_{\{T_i\}_{i=1}^m}(\boldsymbol{X})$. Generally, supposing L_i is the loss function for the task T_i, the total objective function of MTDL can be formulated as $L_{MTDL} = \sum_{i=1}^m \boldsymbol{w}_i \bullet L_i$, where $\boldsymbol{w}_i$ is a weighed term to balance the specific losses of multiple tasks. Furthermore, the network weights in the layers of the deep learning model, denoted as $\boldsymbol{W}$, are updated by optimizing L_{MTDL} via gradient backpropagation.

Similar to the human brain that is capable of multi-tasking, learning with minimal supervision, and generalizing learned skills, all accomplished with high efficiency and low energy cost [11], MTDL brings several advantages. First, it avoids repeated learning of the common-shared features for different tasks, substantially reducing the overall memory consumption. Second, it can learn more generalized features by averaging the inherent noise patterns among various tasks. Third, it can prioritize important features which are difficult to distinguish under the single-task framework. Fourth, it can introduce inductive biases to reduce the overfitting problem, which is superior to conventional regularization methods. In brief, MTDL has the ability to improve the model's efficiency, generalization, and performance with the joint learning of interrelated tasks.

Several kinds of review articles have presented general overviews of multi-task learning [12–14] and its applications in the fields of computer vision [15] and chemoinformatics [16]. However, none of them focused on human medical image computing and analysis. As more and more studies have leveraged MTDL in medical image computing and analysis, the MTDL paradigm is becoming one of the utmost important research topics in the medical community. Although some review articles [2, 17–19] have surveyed deep learning methods applied to medical imaging data, they did not give sufficient information about MTDL. To fill this gap, we present a well-rounded review of the MTDL methods, applications, challenges, and future research trends in medical image computing and analysis.

The arrangement of this paper is as follows. Section 2 summarizes the network architectures or structures of MTDL. Section 3 highlights some MTDL models applied to the image of different human anatomical regions. Section 4 discusses the challenges of MTDL on medical images and pinpoints future research trends.

2. Network architectures

The classification criteria of MTDL network architectures are various. For example. Vandenhende et al. [15] classified the MTDL models into the encoder-focused and decoder-focused architectures; He et al. [20] introduced a hierarchically-fused model and categorized the existing MTDL model into the late or early-branched structure; Zhao et al. [21] classified the existing MTDL works into the structure- and parameter-sharing methods when introducing their proposed flexible and compact multi-task architecture search algorithm. Nevertheless, these taxonomies are too coarse to classify the MTDL architectures constructed for diverse medical image computing and analysis tasks, and inappropriate to classify some MTDL-based studies enrolled in this review. Therefore, we introduce a new taxonomy that focuses on four popular MTDL embodiments, called cascaded, parallel, interacted, and hybrid architectures.

2.1. Cascaded architecture

As illustrated in Fig. 1(a), the output of the previous task is fed into the subnet(s) of the subsequent task(s) without any task-shared layers, i.e., the latter task is dependent on the result of the former task. Mathematically, given an input $\boldsymbol{X}$, the data flow of the cascaded architecture shown in Fig. 1(a) can be formulated as $\{Y_A, Y_B\} = F_B(F_A(\boldsymbol{X}))$, where $Y_A = F_A(\boldsymbol{X})$; $Y_B = F_B(Y_A)$; Y_A and Y_B denote the results of T_A and T_B tasks, shown with the gray and green blocks, respectively; F_A and F_B represent the deep learning subnets for T_A and T_B tasks, shown with the yellow and orange blocks, respectively. This cascaded scheme may be more suitable for solving the task combinations that the subsequent task highly depends on the former task, i.e., the outcome of the former task has a significant influence on the performance of the latter task.

The cascaded MTDL architecture can be trained in an end-to-end way. For example, a segmentation subnet is cascaded to a synthesis subnet for automatic ischemic stroke lesion segmentation [22], where the 'location' is one of key features when describing the clinical impact of a lesion resulting from stroke. Experimental results demonstrate that the synthesized image could improve the segmentation performance compared to the models that directly use perfusion parameter maps, achieving top performance in ISLES 2018 challenge.[3] Alternatively, the cascaded MTDL architecture can be trained in a two-stage [23–25] fashion. For example, Wu et al. [26] propose a two-stage framework based on mesh deep learning (called TS-MDL) for joint tooth labeling and landmark identification on raw intraoral scans. TS-MDL first adopts the iMeshSegNet network to label each tooth. Guided by the segmentation outputs, TS-MDL further selects each tooth's region of interest (ROI) on the original mesh to construct a PointNet network to regress the corresponding landmark heatmaps. Generally, this two-stage framework is applied for works where tasks in MTDL are rather complicated or implemented with limited GPU memory since the end-to-end training may bring higher computation costs.

2.2. Parallel architecture

As illustrated in Fig. 1(b), the task-specific layers are independently parallel to learn the task-specific features for different tasks. Mathematically, given an input $\boldsymbol{X}$, the data flow of the parallel architecture shown in Fig. 1(b) can be formulated as $\{Y_A, Y_B\} = \{F_A(f), F_B(f)\}$, where $f = F_C(\boldsymbol{X})$; $Y_A = F_A(F_C(\boldsymbol{X}))$; $Y_B = F_B(F_C(\boldsymbol{X}))$; Y_A and Y_B denote the results of T_A and T_B tasks, shown with the gray and green blocks, respectively; F_A and F_B represent the deep learning subnets for T_A and T_B tasks, shown with the yellow and orange blocks, respectively; F_C denotes the common task-shared layers shown with blue blocks. This parallel scheme may be more suitable for solving the task combinations that the

[1] http://braintumorsegmentation.org/.
[2] http://adni.loni.usc.edu/.
[3] http://www.isles-challenge.org/.

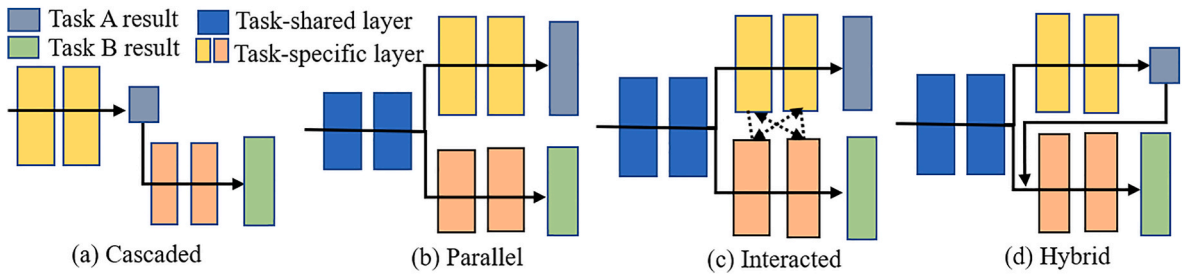

Fig. 1. Four embodiments of network architectures for multi-task deep learning, i.e., (a) cascaded, (b) parallel, (c) interacted, and (d) hybrid.

tasks are related but with different complexities, such as one is the segmentation task and the other is the classification task.

An exemplary parallel architecture is the COVID-19 Multi-Task Network (COMiT-Net) that consists of one task-shared encoder and four task-specific decoders [27]. As shown in Fig. 2(a), the task-shared encoder aims to learn the shared high-dimensional feature representations of the chest X-ray images. The task-specific subnets desire to learn high-level task-specific features to complete the segmentation and classification tasks. The prediction results of eight ablation experiments with different task combinations demonstrate that each task holds its role, i.e., removing either of the three assisting tasks deteriorates the performance. Apart from the last task-shared layer, more previous task-shared layers can be connected to the task-specific subnets. For example, a global average pooling is performed on the output features of each task-shared layer in the shared encoder to form a one-dimensional representation for the glioma classification [28].

2.3. Interacted architecture

As shown in Fig. 1(c), one or more connections exist between task-specific layers, which can dig the deep correlations between related tasks. Mathematically, given an input $\boldsymbol{X}$, the data flow of the interacted architecture illustrated in Fig. 1(c) can be formulated as $\{Y_A, Y_B\} = \{F_{A,B}(f), F_{B,A}(f)\}$, where $f = F_C(\boldsymbol{X})$; $Y_A = F_{A,B}(F_C(\boldsymbol{X}))$; $Y_B = F_{B,A}(F_C(\boldsymbol{X}))$; F_C denotes the common task-shared layers shown with blue blocks; Y_A and Y_B denote the results of T_A and T_B tasks, shown with the gray and green blocks, respectively; $F_{A,B}$ and $F_{B,A}$ represent the interaction between the task-specific layers for T_A and T_B tasks, shown with the black dash lines between the yellow and orange blocks. This interacted scheme may be more suitable for solving the task combinations that the tasks are significantly related so that the in-deep features from the latter deep layers are even able to provide useful auxiliary information to each other.

Typically, the complexity of each task in interacted model is almost similar to each other. As an exemplary interacted architecture shown in Fig. 2(b) [29], the backbone U-Net network delineates the bladder and rectum that are relatively easy to distinguish in CT images. Meanwhile, an attention sub-network is built to hierarchically transfer the backbone features and adaptively learn discriminative representations for the prostate bed segmentation. They are both segmentation tasks with dense outputs. The output of each task can be one label in Ref. [30], where the feature maps from the last convolutional layer of the regression module are concatenated with the feature maps of the classification module. The connections between two subnets in above two studies are one-pass, i.e., the interaction is unidirectional. The connection can be bidirectional, i.e., the feature map of a task-specific subnet can interact with the other subnet and vice versa. For example, the task consistency blocks at multiple levels enable interactions between task-specific layers of the segmentation and contour regression tasks [20].

2.4. Hybrid architecture

Considering the diversity of detailed network structures, there exist some MTDL-based network architectures that not fit nicely into the above three categories. We classify them into the hybrid category. Generally, at least two of the cascaded, parallel, or interacted strategies are leveraged in the hybrid architecture. This hybrid architecture could fully integrate the task-shared and task-specific representations by taking advantage of the pros of the above three architectures, which may be suitable for the complicated task combinations.

Fig. 1(d) shows an exemplary hybrid architecture with both parallel and cascaded schemes, where task A result is obtained based on the task-shared features, and task B result depends on the task-shared features as well as the output of task A. Mathematically, given an input $\boldsymbol{X}$, the data flow of the hybrid architecture presented in Fig. 1(d) can be formulated as $\{Y_A, Y_B\} = \{F_A(f), F_B(f, F_A(f))\}$, where $f = F_C(\boldsymbol{X})$; $Y_A = F_A(f)$; $Y_B = F_B(f, Y_A)$; F_C denotes the common task-shared layers shown with blue blocks; Y_A and Y_B denote the results of T_A and T_B tasks, shown with the gray and green blocks, respectively; F_A and F_B represent the task-specific layers for T_A and T_B tasks, shown with the black dash lines between the

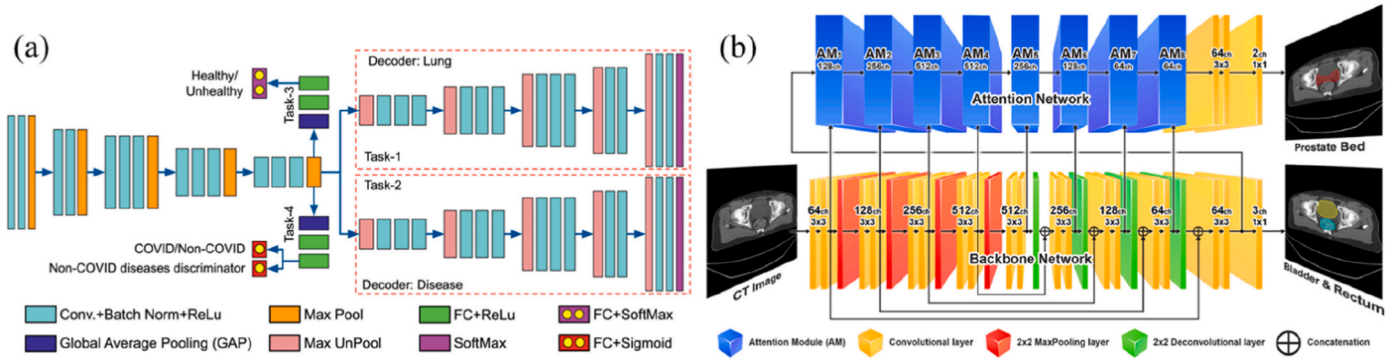

Fig. 2. (a) A parallel multi-task deep learning architecture contains four subnets for two classification tasks and two tasks of lung and disease region segmentation [27]. (b) An interacted multi-task deep learning architecture where the prostate bed segmentation subnet is based on the features from the bladder and rectum segmentation subnet [29].

yellow and orange blocks.

The hybrid model can be constructed by integrating parallel and interacted architectures [31,32]. For example, the hybrid MTDL architecture shown in Ref. [32] incorporates both the parallel and interacted structures. Specifically, the encoder is shared by the two decoders, which are trained to segment either bright or red lesions. The generalization block is trained to predict whether the input contains any lesions. The classification subnet is parallel to the two segmentation subnets, while the two segmentation subnets are interacted with each other. Besides, some complex hybrid models can even exploit cascaded, parallel, and interacted strategies under one framework [33–35]. For example, the hybrid model in Ref. [35] consists of three subnets, where a coarse segmentation network (coarse-SN) generates coarse lesion masks that provide a prior bootstrapping for a mask-guided classification network (mask-CN) to help it locate and classify skin lesions accurately, and the lesion localization maps produced by mask-CN are then fed into an enhanced segmentation network (enhanced-SN) to help it delineate lesion region. The classification network is cascaded to the coarse segmentation network and interacts with the enhanced segmentation network, while the coarse segmentation network and the enhanced segmentation network are parallel to each other. In this way, both segmentation and classification networks mutually transfer knowledge between each other and facilitate each other in a bootstrapping way.

3. Highlighted case studies

MTDL-based methods have been applied to medical imaging of various human anatomical regions. This section gives an overview and highlights some advances in diverse clinical practices.

3.1. Brain

Neuroimaging plays a paramount role in the diagnosis of many life-threatening brain diseases, such as glioblastoma [36], intracranial hemorrhage [37], ischemic stroke [38], schizophrenia [39], and Alzheimer's disease [40]. Table 1 lists some representative works that exploit MTDL to enhance the joint learning performance on brain image. Most MTDL models directly take the brain images as input. While four MTDL models perform on the graph-based adjacency matrix constructed from fMRI. The adjacency matrix may refer to the similarity matrix [41] or functional network connectivity [39] that is specific in neuroscience.

As shown in Table 1, more than half of the papers focus on image segmentation of multiple brain regions-of-interest (ROIs) [62,63], or combining it with image regression [38,40], classification [64], synthesis [46,42], reconstruction [54], registration [53], and other relevant tasks [52]. Specifically, the hippocampal segmentation task can work as the prerequisite step for the disease classification task, and a 3D hybrid MTDL model is designed by inputting the features from the segmentation subnet and the corresponding segmentation results into the classification subnet to identify AD patients [64]; Whereas image synthesis task can work as the prerequisite step for the brain tumor segmentation, and a cascaded MTDL model is constructed by feeding the synthesized images with high tissue contrast into the segmentation subnet to detect tumor and tumor cores, jointly trained in an end-to-end fashion [42]. Compared with the two-stage counterparts requiring separate training steps, i.e., minimizing synthesis loss to produce images with high tissue contrast and then optimizing segmentation loss, the end-to-end architecture can yield higher accuracy.

Some papers have mainly concentrated on image-level classification [58,55] and regression [36,56,57]. For example, a parallel MTDL model combines point-detection and angle-detection into a unified structure, where the point-detection is used to localize the AC and PC points, and the angle-detection is used to determine the angulation of the line connecting these points [49]; An interacted MTDL architecture is designed by concatenating high-level features obtained from the tumor genotype classification subnet with features used for overall survival time regression [36]. Four papers have incorporated classification and synthesis tasks into one MTDL network [39,44,45,47]. For instance, a cascaded MTDL model is established by inputting the enhanced 3T MRI image generated based on the corresponding 1.5T MRI into Alzheimer's disease classification subnet [47].

Table 1
Some representative literatures using MTDL for brain image analysis.

Architecture	Ref.	Tasks
Cascaded	[42]	Target-modal synthesis and tumor/brain structure segmentation[a]
	[22]	DWI image synthesis and lesion segmentation
	[43]	Single-contrast and multi-contrast super-resolution reconstruction
	[44]	MR sequences synthesis and image-contrast classification
	[39]	FNC generation and disease classification
3D Cascaded	[45]	Image generation and disease classification
	[46]	Image synthesis at two future time-points
	[47]	Image synthesis and disease classification[b]
Parallel	[41]	Region-wise segmentation of four eloquent cortex areas
	[48]	Target-modal synthesis and image edge prediction
	[49]	Landmark point center localization and angle detection
	[37]	ICH region segmentation and foreground/background reconstruction
3D Parallel	[28]	Tumor segmentation and glioma subtyping
	[50]	Glioma segmentation and IDH genotyping
	[51]	Tumor segmentation and image reconstruction
	[38]	Early infarct segmentation and ASPECTS scoring
	[40]	Hippocampus segmentation and MMSE score regression
	[52]	Tumor segmentation and distance map estimation
	[53]	Image registration and brain anatomical segmentation
	[54]	Tumor/WMH segmentation and foreground & background reconstruction
	[55]	Classifications of AD/HC and pMCI/sMCI [c]
	[56]	Disease classification and four clinical scores regression
	[57]	Three clinical scores prediction
	[58]	Three classifications of glioma
Interacted	[59]	Multi-modal image reconstruction and target-modal synthesis[d]
3D Interacted	[36]	Genotype prediction and overall survival time prediction
Hybrid	[60]	Three classifications, central and peripheral brain segmentation
	[61]	Target-modal synthesis, source-modal reconstruction, and tumor segmentation[e]

[a] https://github.com/hamghalam/HTC-segmentation.
[b] https://github.com/vkola-lab/azrt2020.
[c] https://github.com/simeon-spasov/MCI.
[d] https://github.com/taozh2017/HiNet.
[e] https://github.com/devavratTomar/sasan.

3.2. Eye

Benefiting from the convenient acquisition of retinal images, a wide variety of MTDL models have been constructed to accomplish multiple tasks, such as DR grading [65], lesion segmentation [66], and vessel segmentation [67]. For example, considering the correlation between bright and red lesions, i.e., the presence of one type of lesion is a strong sign for the potential presence of the other type, Playout et al. [68] proposed a hybrid MTDL architecture to predict whether the input contains any lesions and to segment either bright or red lesions, where the exchange layer aimed to communicate information between the two segmentation branches. Based on the findings that the super-resolved input images can improve the performance of DR grading and lesion segmentation, and the lesion segmentation regions of fundus images are highly consistent with pathological regions for DR grading, a cascaded structure was proposed to simultaneously process the low-level task of image super-resolution (ISR), the mid-level task of lesion segmentation and the high-level task of DR grading, by feeding the outputs of ISR and segmentation subnets into the DR grading subnet [69].

Table 2 lists some representative MTDL-based works for retinal image analysis and computing. Most performed for DR detection on

Table 2
Some representative literatures using MTDL for retinal image analysis.

Architecture	Ref.	Tasks
Cascaded	[73]	SLO image generation and image classification
	[74]	Image generation and normal/abnormal classification
Parallel	[75]	Optic disc and cup segmentation, contour detection, distance map estimation
	[70]	Multi-view instance discrimination and rotation prediction
	[76]	Thick and thin vessel segmentations
	[72]	DR, AMD, GON, melanoma and normal classification, 320 fine-grained disease sub-categories prediction, textual diagnosis generation[a]
	[65]	DR grading and AMD grading
	[77]	DR grading and lesion segmentation
Interacted	[71]	DR related features regression and DR severity classification
	[30]	Visual field measurement regression and Glaucoma classification
	[67]	Vessel type classification and vessel similarity estimation
	[78]	Image generation and classification
	[66]	DR grading and lesion segmentation [b]
	[79]	Diabetic retinopathy (DR) and diabetic macular edema (DME) classifications[c]
Hybrid	[80]	Retinal full vessel and artery/vein segmentations
	[32]	Lesion detection and red and bright lesion segmentations
	[69]	Image super-resolution (ISR), lesion segmentation, and diabetic retinopathy (DR) grading

[a] https://github.com/SahilC/model-kd-disease-recognition.
[b] https://csyizhou.github.io/FGADR/.
[c] https://github.com/xmengli999/CANet.

color fundus images, mainly because of the early public available dataset (listed in Ref. [66]) and the release of Kaggle Competition in 2015. Recently, some large datasets for multiple tasks have been collected, annotated, and even released [66,70]. For example, to improve the interpretability of diagnostic results, a total of 89917 digital funds images have been collected and annotated with twelve DR-related features and DR severity [71]. And an interacted MTDL model has been established to explore their causal relationship by incorporating the properties of DR-related features into DR severity diagnosis. Chelaramani et al. [72] have collected 7212 labeled and 35854 unlabeled fundus images and trained a semi-supervised MTDL model to classify the disease category, predict fine-grained disease sub-categories, and generate a textual diagnosis. These studies may open up many interesting avenues of retinal image analysis.

3.3. Chest

Table 3 lists some representative MTDL-based works implemented on chest imaging. Most studies have adopted the parallel architecture to simultaneously fulfill multiple tasks, such as segmentation [9,81], classification [82], localization [83], and characterization [84]. For example, a 3D parallel MTDL network, namely *Deep Profiler*, can predict time-to-event treatment outcomes and approximate classical radiomics based on the pre-therapy lung CT images [85]. The cascaded [9,84], interacted, and hybrid MTDL architectures have also been exploited. For example, an interacted MTDL is established by interacting the features from the classification branch with the features in the encoder part of the segmentation branch and then feeding the interacted features into the decoder part of the segmentation branch to improve the performance of lesion localization [82]; A hybrid MTDL model proposed for COVID-19 diagnosis and severity assessment incorporates cascaded and parallel structures by inputting the shared stacking 2D feature maps extracted with U-Net backbone into two parallel subnets of COVID-19 classification and lesion segmentation and meanwhile performing the regression of severity score based on the lesion segmentation results [86].

The outbreak of COVID-19 and the publicly available datasets, e.g.,

Table 3
Some representative literature using MTDL for thoracic imaging analysis.

Architecture	Ref.	Tasks
Cascaded	[84]	Lung nodule malignancy classification and nodule features characterization
3D Cascaded	[9]	Lung nodule segmentation and patient-level malignancy classification
Parallel	[87]	Lung and nodule segmentations
	[88]	Nodule classification and image reconstruction
	[81]	Thoracic organ segmentation and multi-label classification of organs[a]
	[89]	Lung nodule detection and segmentation
	[90]	Line segmentation and tip detection of PICC
	[91]	Instance level detection and segmentation
	[92]	Lesion segmentation, COVID-19 diagnosis, and CT image reconstruction[b,c]
	[93]	Diagnosis and severity quantification of COVID-19
	[94]	Lobe segmentation and multi-instance severity classification [d]
3D Parallel	[95]	Deformable registration and nodule classification
	[96]	Deformation vector fields generation
	[85]	Radiomics estimation and therapy outcome prediction
	[97]	Tumor segmentation and nodule classification
	[98]	Pulmonary lobes and lobe border segmentation
	[99]	False positive nodule reduction and segmentation
	[100]	Lung nodule detection and malignancy classification
	[101]	Orientation field regression, heatmap regression, and tree structure segmentation for lung airway landmark detection
Interacted	[102]	Lung node benign/malignant classification and attribute score regression
	[82]	Lung lesion segmentation and COVID-19 infected or uninfected classification [e]
Hybrid	[103]	Multi-scale segmentation and classification
	[34]	Lung nodule segmentation, malignancy classification, and medical features predictions
	[86]	Identification of COVID-19 and severity quantification[f]
3D Hybrid	[104]	Lung nodule segmentation, attributes, and malignancy prediction
	[105]	Lung nodule detection and segmentation[g]
	[33]	Lung lesion segmentation and disease classification[h]
	[83]	Lung segmentation, nodule center identification and size regression
	[106]	Overall survival risk prediction, tumor stage prediction and node stage predictions
	[107]	Image reconstruction, tumor segmentation, classification (esophageal vs lung cancer), and a multi-scale outcome prediction

[a] https://github.com/ithet1007/MTL-SegTHOR.
[b] https://github.com/UCSD-AI4H/COVID-CT/.
[c] http://medicalsegmentation.com/covid19/.
[d] https://github.com/KeleiHe/M2UNet.
[e] https://github.com/yuhuan-wu/JCS.
[f] https://github.com/neuro-ml/COVID-19-Triage/tree/master/covid_triage.
[g] https://github.com/uci-cbcl/NoduleNet.
[h] https://github.com/XiaofeiWang2018/DeepSC-COVID.

LUNA16[4] have incentivized the development of MTDL-based methods on chest X-ray or CT images. Some existing COVID-19 databases have been summarized in Ref. [33]. Moreover, many studies have established in-house datasets for their specific clinical applications [85,90,96,97]. For example, to help nurses automatically and timely identify the peripherally inserted central catheter (PICC) position in X-ray images, Yu et al. have collected chest radiographs from 326 patients with visible PICC, only including anteroposterior projection viewpoint in the dataset [90].

3.4. Cardiac

Table 4 lists some representative MTDL-based models performed on cardiac imaging for the joint learning of segmentation [108], motion

[4] https://luna16.grand-challenge.org/.

Table 4
Some representative literatures using MTDL for cardiovascular imaging analysis.

Architecture	Ref.	Tasks
Cascaded	[118]	Segmentations of left atrium and atrial scars
	[119]	LV contour segmentation and full LV quantification
	[115]	LV segmentation and LV quantification
3D Cascaded	[120]	Scene registration and seven cardiac structures segmentation
Parallel	[117]	Full LV quantification and cardiac phase identification
	[114]	Calcification segmentation and artery-specific calcification quantification
	[121]	Full LV quantification and cardiac phase classification
	[122]	LV segmentation and parameters estimation
	[10]	Bi-ventricle segmentation and direct area estimation
	[123]	Cardiac segmentation and motion estimation
	[124]	Colorize, determine rotation and localize tasks
	[125]	Left atrial segmentation and pre/post-ablation classification
	[126]	Survival prediction and surface mesh reconstruction [a]
	[127]	LV segmentation and landmark localization
3D Parallel	[109]	LV segmentation and motion tracking in 4D echocardiography
	[112]	Centerline distance map and endpoint confidence map estimations
	[128]	Classifications of coronary artery plaque and stenosis
Interacted	[108]	Infarction area segmentation and quantification
	[116]	LV segmentation and full LV quantification
	[129]	LV segmentation and full LV quantification
	[110]	Perfusion parameter estimation and ischemia classification
Hybrid	[130]	Multitype cardiac indices estimation, cardiac segmentation, and image reconstruction
3D Hybrid	[113]	LV cavity and Myo segmentation, full LV quantification, and phase classification [b]

[a] https://github.com/UK-Digital-Heart-Project/4Dsurvival.
[b] https://github.com/sulaimanvesal/CardiacQuanNet.

estimation or tracking [109], parameter quantification [110], and other related tasks [111,68]. Most are CNN-based MTDL [112–114], while a few combine CNN with long short-term units (LSTM) or recurrent neural network (RNN) to take advantage of both the spatial and temporal dynamics of image slices [115–117]. For example, the Bi-ResLSTM units that can capture spatial-temporal behavior patterns along the cardiac cycle are implemented to carry out the simultaneous multi-view segmentation and multidimensional quantification of LVs from paired echo sequences [116]; The MTDL model in Ref. [117] first obtains expressive and robust cardiac representations with a CNN structure and then models the temporal dynamics of cardiac sequences with two parallel RNN modules to estimate the three types of LV indices and the cardiac phase, respectively. Each network architecture has been utilized for cardiac image computing and analysis. For example, a Multi-view Weighted Fusion Attention (MMWFAnet) model in the parallel MTDL can extract discriminative feature representations from multiple views of non-contrast CT scans, which are then fed into a segmentation subnet and a regression subnet to obtain accurate segmentation and quantification of artery-specific calcification simultaneously [114].

Many MTDL models involve the full LV quantification tasks, i.e., estimation of cavity and myocardium areas, dimension of LV cavity, and regional wall thicknesses (RWTs). The open Left Ventricle Full Quantification Challenge (LVQuan18)[5] partially drives this trend. At the same time, other public datasets also promote the exploit of MTDL models, such as the UK Digital Heart Project Dataset [68], Automated Cardiac Diagnosis Challenge (ACDC) [7] [122], Atrial Segmentation Challenge2018[6] [111], and 3D Strain Assessment in Ultrasound (STRAUS) [109]. In addition, several specific datasets have been built for some particular tasks. For example, the retrospectively collected CT myocardial perfusion images of 232 in-house patients are harnessed to train a Spatio-temporal Multitask Network Cascade (ST-MNC) module to predict various perfusion parameters and classify myocardial ischemic regions simultaneously [110]; And CMR images (20 frames per cycle) of 302 patients with pulmonary hypertension have been collected for survival prediction, with the auxiliary task of reconstructing 20 frame-wise cardiac motion meshes [126].

[5] https://lvquan18.github.io/.
[6] https://www.creatis.insa-lyon.fr/Challenge/acdc/.
[7] https://github.com/cherise215/atria_segmentation_2018/.

3.5. Abdomen

Table 5 lists some representative MTDL-based models performed on the abdominal imaging. In terms of task combination, two-thirds incorporate the segmentation task [131], while others are concerned about the tasks such as classification [132,133], registration [134], landmark detection [135], and survival prediction [136,137]. In terms of network architecture, half leverage the parallel architecture. For example, a 3D CNN-based parallel MTDL model is constructed to segment gastric tumors and classify lymph nodes simultaneously with two task-specific subnets based on the task-shared features extracted from the scale-aware and task-aware attention-guided learning module [138]. A parallel MTDL model with Contrast-Enhanced Convolutional

Table 5
Some representative literatures using MTDL on abdominal image.

Architecture	Ref.	Tasks
Cascaded	[142]	Abdomen organ segmentation and domain translation (generation) [a] [b]
	[141]	Liver segmentation and domain translation [c]
	[140]	Kidney segmentation and domain translation
	[139]	Splenomegaly segmentation and domain translation[d]
	[145]	Probe localization and image inpainting
3D Cascaded	[146]	Multi-modal image synthesis, registration, and liver segmentation.
Parallel	[147]	Eroded and dilated mask generation; soft and hard label for pelvic organ segmentation
	[148]	Prostate, bladder and rectum segmentation and prostate boundary regression
	[149]	Kidney tumor detection, segmentation, and quantification
	[136]	Six DL-based survival prediction and radiomics feature reconstruction
	[133]	11 phases and 44 steps classification[e]
	[132]	Image position and direction classification
	[135]	View classification and landmark detection
	[150]	Subcutaneous and visceral fat maps prediction [f]
3D Parallel	[151]	Liver vessel extraction, centeredness score determination, and connectivity estimation between center-voxels
	[137]	Resection/margin status and overall survival prediction
	[152]	Liver segmentation and kidney segmentation
	[131]	Tumor proximal region segmentation and distal region segmentation
	[153]	Pancreas segmentation and its skeleton extraction
	[138]	Tumor segmentation and lymph node classification [g]
Interacted	[143]	Image classification and lesion segmentation
	[29]	Prostate bed (PB) segmentation, and bladder & rectum segmentation [h]
3D Interacted	[154]	Tumor segmentation and pre/post-CRT classification[i]
	[20]	Gland segmentation and gland contour delineation
	[155]	Organ segmentation and image registration[j]
Hybrid	[134]	Four registration tasks
	[144]	Auxiliary characteristics regression and disease classification

[a] https://chaos.grand-challenge.org/.
[b] https://github.com/harveerar/PSIGAN.
[c] https://github.com/bbbbbbzhou/APA2Seg-Net.
[d] https://github.com/MASILab/SynSeg-Net.
[e] https://github.com/CAMMApublic/MTMS-TCN-Phase-Step-Bypass.
[f] https://www.cancerimagingarchive.net/.
[g] https://github.com/infinite-tao/MA-MTLN.
[h] https://github.com/superxuang/amta-net.
[i] https://github.com/Heng14/3D_RP-Net.
[j] https://github.com/moelmahdy/JRS-MTL.

LSTM (CE-ConvLSTM) module is designed to accomplish two tasks of survival outcome and margin prediction [137]. Several MTDL models utilize the cascaded architecture, primarily by cascading a generation subnet with a segmentation subnet, such as for the segmentation of splenomegaly [139], kidney [140], liver [141], and multi-organ in the abdomen [142]. The interacted architecture is also exploited. For example, a deep synergistic interaction network (DSI-Net) consists of a backbone network and three task-specific subnets of classification branch (C-Branch), coarse segmentation branch (CS-Branch), and fine segmentation branch (FS-Branch), where the feature maps from CS-Branch are fed into C-Branch and FS-Branch, and the prototype center provided by C-Branch is fed into FS-Branch to improve the performance of joint classification and segmentation on the wireless capsule endoscope (WCE) images [143]. Moreover, two hybrid MTDL models have been constructed to register images with large deformation [134], and diagnose advanced gastric cancer [144].

The majority of MTDL models perform on privately collected abdominal images, mainly including image modalities of CT scan [147], multiparametric MRIs [154], ultrasound [135], and even video recordings of laparoscopic gastric bypass procedures [133]. It is worth noting that some research efforts adopted MTDL to accomplish some infrequent tasks, such as treatment response prediction [154], prediction of fat distribution [150], probe localization [145], and resection margin estimation [156]. For example, a multitask multi-stage temporal convolutional network (MTMS-TCN) implemented on the in-house video cohort consisting of 40 surgical procedures can jointly predict two correlated surgical activities, i.e., phases and steps, which may provide a novel solution for evaluation of the execution of surgical procedures [133].

3.6. Musculoskeletal

Table 6 lists some representative MTDL-based models using the musculoskeletal images. The tasks involve bone segmentation [157], landmark localization [158], image content classification [159,160] and other related tasks [56]. Regarding the network architecture, the popular is parallel, followed by interacted, hybrid, and cascaded ones. Take a parallel architecture as an example, a sequential conditional reinforcement learning network [161] is proposed to tackle the simultaneous detection and segmentation of vertebral body (VB) from MR spine images, where a subnet named fully-connected residual neural network learns rich global context information of the VB including both the detailed low-level features and the abstracted high-level features to detect the accurate bounding-box of the VB; and another subnet named Y-shaped Network learns comprehensive detailed texture information of VB including multi-scale, coarse-to-fine features to segment the boundary of VB.

From Table 6, we can observe that both the task combination and the imaging modality are various. This phenomenon may be due to the diversity of musculoskeletal diseases or applications and the limited public datasets. In fact, most studies have collected and labeled the in-house datasets for their specific tasks [164,165,172]. For example, Schacky et al. [166] have regarded the five hip osteoarthritis features assessments as multiple classification tasks and trained the encoder-focused parallel MTDL model to grade these features simultaneously.

Table 6
Some representative literatures using MTDL for musculoskeletal image analysis.

Architecture	Ref.	Tasks
Cascaded	[162]	Vertebra segmentation and classification
	[163]	Volume projection imaging (VPI) image restoration and spine segmentation
Parallel	[161]	Vertebral body (VB) detection, segmentation, and classification
	[164]	Femur region and boundary identification
	[165]	Bone shadow enhancement (BSE) and horizontal bone interval mask (HBIM) prediction
	[158]	Anatomical structure detection, segmentation, and landmark localization
	[166]	Five classifications of hip osteoarthritis features per joint [a]
	[157]	Bone segmentation, line and landmark localization
	[167]	Implant brand and treatment stage classification
	[159]	Gender and age group classifications
Interacted	[168]	Vertebral segmentation and landmark localization
	[169]	Multi-scale segmentations of intervertebral discs
	[170]	Intervertebral disc, vertebra, and neural foramen detection, segmentation, and classification
Hybrid	[171]	Vertebral localization, identification, and segmentation

[a] https://github.com/Rad-190925/Code.

3.7. Digital pathology and microscopy

Some studies have applied MTDL on gigapixel whole-slide images (WSI) and tissue slide images to accomplish the tasks such as detection, segmentation, and classification. Table 7 lists some representative papers. All MTDL models are 2D-based architectures due to the attribute of pathological and microscopical images. However, the image size can be over $10^5 \times 10^5$ pixels, with exceptionally high resolution. Considering GPU memory consumption, it is almost unfeasible to directly input the whole image into the MTDL model. Most studies have divided the entire image into small patches or instances and separately fed them into the model. The first parallel MTDL model aims to carry out the tasks of mass detection and classification based on breast cancer histopathology images [173]. Since then, MTDL has been showing its prominent performance. For example, Chen et al. [174] have proposed a parallel MTDL model to detect the boundary of a gland as well as perform normal pixel segmentation, ranking first in the 2015 MICCAI Gland Segmentation Challenge by a large margin. Wang et al. [31] have presented a novel

Table 7
Some representative literatures using MTDL for digital pathological and microscopical image analysis.

Architecture	Ref.	Tasks
Cascaded	[175]	Instance-level segmentation and diagnostic classification [a]
	[176]	Tumor classification and segmentation
Parallel	[173]	Malignant/benign and image magnification level classifications
	[177]	Inner and outer distance maps estimation [b]
	[174, 178]	Gland segmentation and contour detection
	[179]	Lesion classification and segmentation
	[180]	Image classification and retrieval
	[181]	Multi-class recognition task and verification task of image pairs
	[182]	Classify sperm's head, vacuole, and acrosome as either normal or abnormal[c]
	[183]	Cell segmentation and marker prediction[d]
	[184]	Gland segmentation, lumen segmentation, nuclear segmentation and tissue type classification
	[185]	Cancer region detection and subtyping
	[186]	Cancer classification and gland segmentation
Interacted	[187]	Primary or metastatic tumor classification and organ site classification[e]
	[188]	Chromosome joint detection, chromosome type and polarity classification
	[189]	Malignant/benign classification and gland segmentation
Hybrid	[190]	Multi-instance localization and image classification
	[31]	Hepatocellular carcinoma segmentation and classification
	[191]	2-class and 5-class classifications and a manual features fitting task

[a] https://sacmehta.github.io/YNet/.
[b] http://www.cs.bilkent.edu.tr/~gunduz/downloads/DeepDistance.
[c] https://github.com/amirabbasii/The-Blessing-of-Deep-Transferand-Multi-task-Learning-on-Sperm-Abnormality-Detection.
[d] https://github.com/291498346/nas_cellseg.
[e] https://github.com/mahmoodlab/TOAD.

hybrid MTDL model to perform the hepatocellular carcinoma (HCC) segmentation and classification with three task-specific branches, achieving second place in the MICCAI 2019 Pathology AI Platform (PAIP) challenge.[8]

3.8. Miscellaneous

Apart from the above body sites, computer-aided automatic computing and analysis are also carried out on the medical imaging of other anatomical regions or organs. For example, teeth [25,192–195], as the only masticatory organ in the human digestive system, are responsible for food chewing, auxiliary pronunciation, and facial morphology development. Recently, many DL-based methods have been widely applied in the dental field [196–198] and are also a recent research hotspot. Table 8 lists some representative MTDL-based studies focused on the image analysis of teeth and other body sites. For example, considering the inherent association of landmarks and bone segmentation, i.e., landmarks generally lie on the boundaries of segmented bone regions, the first 3D hybrid MTDL model is established to jointly segment dental cone-beam computed tomography (CBCT) images into midface and mandible and digitize 15 anatomical landmarks [199]. All MTDL models for skin lesion detection, segmentation, and classification have been performed on the 2D skin dermoscopic images, which can uncover the detailed morphological and visual properties of pigmented lesions. Public datasets such as Digital Database for Screening Mammography (DDSM) and INbreast have promoted the development of MTDL to skin images. Kawahara et al. [200] have collected private skin dermoscopic and clinical images and proposed the first parallel MTDL model to concurrently predict the entire 7-point criteria and the diagnosis in a single optimization. Recently, MTDL-based fetal head image analysis has been emerging for assessing brain development and detecting abnormalities [201,202]. As for the breast image analysis, two-thirds of the papers have utilized the public datasets, such as International Skin Imaging Collaboration (ISIC) 2017 Challenge.[9] Almost all studies have established their own task-specific datasets for the image analysis of fetal head [201], tongue [203], and parotid [142]. In contrast, Namburete et al. [204] have innovatively trained a 2D–CNN–based parallel MTDL model on a public dataset to simultaneously derive fetal brain orientation, eye localization, and brain masking. Compared with independent training for each task, the joint learning of these closely related tasks not only saves on training time and memory requirements, but also can improve the performance of tasks with unbalanced data labels. Xu et al. [203] have collected 1858 tongue images by a specialized device with a high-end industrial CCD camera to train a cascaded MTDL model that can simultaneously address the interrelated tasks of tongue image segmentation and classification. These attempts may motivate more new applications of MTDL in medical image computing and analysis.

Table 8
Representative literatures using MTDL for the image analysis of other human body regions.

Body Region	Architecture	Ref.	Tasks
Tooth	Cascaded	[24]	Dental crown surface reconstruction (coarse-to-fine)
		[26]	Tooth segmentation and landmark localization/regression
	Parallel	[205]	Detection and classification of dental diseases
	3D Parallel	[206]	Multi-branch tooth segmentation (tooth region and surface prediction)
		[207]	Prediction of centroid and skeleton offsets; Tooth segmentation, boundary detection, and landmark localization
	3D Hybrid	[208]	Tooth detection, tooth segmentation and tooth ID prediction
		[209]	Craniomaxillofacial landmark localization (coarse-to-fine)
		[210]	Tooth segmentation, ID classification, and alveolar bone segmentation
		[199]	Displacement map estimation, bone segmentation, and landmark digitization
		[211]	Mandible segmentation, meta-level, and six landmarks localization
Skin	Parallel	[212]	Skin lesion detection, classification, and segmentation
		[200]	7-point melanoma checklist criteria classification and skin lesion diagnosis
	Interacted	[213]	Skin lesion classification and segmentation
		[214]	Skin lesion segmentation and edge prediction
	Hybrid	[35]	skin lesion segmentation and classification
Breast	Parallel	[215]	Abnormal detection and image classification
		[216]	Abnormal detection and segmentation
		[217]	Mass classification and segmentation
	3D Parallel	[218]	Tumor classification and segmentation
	Interacted	[219]	Mass detection, segmentation, and classification [a]
		[220]	Tumor classification and segmentation
Fetal head	Parallel	[204]	Image classifications, eye socket localization and fetal brain extraction
		[201]	HC ellipse segmentation and ellipse parameter estimation
		[221]	Anatomical structures localization and image quality assessment
		[202]	Image quality assessment and fetal brain extraction
Head	Parallel	[222]	Classification, segmentation, reconstruction
Tongue	Hybrid	[203]	Tongue segmentation and tongue coating classification
Carotid and thyroid	Cascaded	[23]	High-quality image reconstruction (coarse-to-fine)
Nasopharynx	3D Hybrid	[223]	Survival prediction and tumor segmentation
Multi- sites	Hybrid	[224]	Multi-organ classification, detection, and segmentation
	Cascaded	[225]	Synthesis and site classification
	Parallel	[226]	Organ segmentation and modality classification
	Interacted	[227]	Abnormal lymph node detection

[a] https://github.com/matterport/Mask_RCNN.

4. Discussion and conclusion

In this review, we first introduce the motivation for surveying MTDL-based medical image analysis and computing. Second, we summarize four popular MTDL network architectures and explain the related network activities with mathematical formulas and typical exemplary embodiments. Then, we list some representative studies to overview the research actuality of MTDL performed on the images of various body sites, including the tasks, network architectures, and image datasets in the specific fields, such as brain, eye, and chest. From the above studies, we can observe that MTDL is an effective and valuable paradigm for a wide range of medical image analysis and computing tasks. However, several unique challenges need to be overcome for future clinical practices. In this section, we will discuss current challenges and potential research directions from the following five aspects.

4.1. Architecture design

This review concentrates on four popular MTDL architectures of parallel, cascaded, interacted, and hybrid. Tables 1–8 show that each MTDL architecture has been applied to the medical imaging of diverse

[8] https://paip2019.grand-challenge.org/.
[9] https://www.isic-archive.com/.

human anatomical regions. Thus, the architecture selection is independent of different imaging modalities of various organs, but may be determined by the complexity of tasks and their relationships.

To be specific, the cascaded architecture may be more suitable for solving the task combinations that the subsequent task highly depends on the former task, i.e., the outcome of the former task has a significant influence on the performance of the latter task. The parallel architecture may be more applicable for solving the task combinations in that the tasks are related but have different complexities, e.g., one is the segmentation task and the other is the classification task. The interacted architecture may be more appropriate for solving the task combinations where the tasks are with similar complexity and significantly related so that the in-deep features from the latter deep layers can even provide useful auxiliary information to each other. The hybrid architecture could fully integrate the task-shared and task-specific representations by taking advantage of the pros of the above three architectures, which may be suitable for the more challenging tasks or task combinations.

In practice, the detailed structures may be diverse, since the integration of common-shared and task-specific layers can be various, such as three variants employed in Refs. [77,228]. For example, Foo et al. [77] have explored three structures to perform DR grading and lesion segmentation, respectively named as variant A (multiple outputs at the decoder part of U-Net, parallel architecture with shared U-net), variant B (the segmentation output is the input of grading subnet, cascaded architecture), variant C (parallel architecture with shared encoder part of U-net). The experimental results show that variant C achieves the best performance in both segmentation and classification tasks. Thus, it needs to carefully and elaborately choose or construct specific architectures, e.g., decide which layers to share or branch out as task-specific layers, based on the given tasks and training data. However, such experimental or handcrafted features are cumbersome to obtain, and the model construction is less efficient.

Some studies have harnessed the neural architecture search (NAS) technique to automatically construct deep neural networks to achieve optimal performances in the given tasks [2]. In the field of computer vision, Gao et al. [229] have incorporated NAS into general-purpose multi-task learning and propose to disentangle multi-task networks into single-task backbones under a hierarchical and layer-wise feature sharing scheme, which is different from typical NAS methods that define search spaces according to task characteristics. Zhao et al. [21] have proposed a novel NAS approach to discover flexible and compact MTDL architectures that can adaptively share structure (parameters) in processing different levels of task relatedness, resulting in further performance improvement. Recently, a NAS-based solution has been proposed to identify optimal networks for joint cell segmentation and marker identification in time-lapse microscopy images, demonstrating the potential of NAS in the construction of MTDL models [183].

4.2. Task selection and relationship modeling

MTDL can carry out various task combinations on diverse medical imaging modalities. Some tasks are related and complementary to each other, while others may be unrelated or even competing. The joint learning performance may not be improved if the tasks are not appropriately combined. Currently, most papers have claimed that there exist task relationships in their constructed MTDL models and then conducted ablation experiments to verify their postulated hypotheses. It would be more efficient to identify and measure related tasks in more objective way. But how to quantify, depict, or determine the relatedness of multiple tasks is still an open question.

Several researches may provide some insights to model task relationships [69,114,230]. In detail, Wang et al. have performed task correlation analyses to demonstrate that the tasks of image super-resolution, lesion segmentation, and diabetic retinopathy grading are closely related [69]; Zhang et al. [114] have introduced task-guided constraint into the MTDL model to learn task dependencies between the segmentation and quantification of coronary artery calcification, which can model the correlation of these two tasks more effectively by exerting segmentation-guided constraint on task-aware feature learning. In the field of computer vision, Fifty et al. [230] have proposed an efficient approach for the identification of task grouping, by co-training all tasks together and quantifying the effect to which one task's gradient would affect another task's loss, which may be effective in task correlation detection in future MDTL-based medical image computing and analysis.

4.3. Multi-task loss optimization

As the loss function governs multi-task deep learning, the design and balance of task losses are therefore critical for joint learning performance. Considering there are two tasks in MTDL with corresponding losses of L_1 and L_2, the final objective function is formulated as $L = w_1L_1 + w_2L_2$, where w is a hyper-parameter to balance the task-specific losses of L_1 and L_2. Improper settings of w may induce task dominance and significantly reduce the overall performance. Existing research works have tried to balance the corresponding losses with different hyper-parameters (such $k = w_1/w_2 = 0.05, 0.1, \ldots, 8$ [231]) and selected the ones that perform best [20,33,143]. Such a manual selection approach requires enormous time and memory costs. The selected hyper-parameters may not be optimal because of the limited search space (maybe best performance achieved when $k = 25$). Moreover, considering data variances across different datasets, the hyper-parameters of the learning weights selected for one cohort may not be suitable for other datasets. Thus, more adaptive optimization strategies need to be designed to balance specific losses of multiple tasks.

Recently, some studies have adopted the adaptive strategies [50,138,232] for more efficient multi-task optimization. For example, Zhang et al. [138] have used an uncertainty weighting strategy to optimize multi-task loss for the joint segmentation of gastric tumors and classification of lymph nodes. Specifically, the total loss function is defined as $L = \frac{1}{2\sigma_1^2}L_{SDS} + \frac{1}{2\sigma_2^2}L_{class} + log\,\sigma_1\sigma_2$, where L_{SDS} and L_{class} denote the segmentation loss and the classification loss, respectively. σ_1 and σ_2 indicate uncertain weights and are obtained through network learning. In practical applications, the parameters σ_1 and σ_2 are first initialized as two tensors with value of 1, and then they are iteratively updated during training phase. In comparison, Bao et al. [233] have proposed a novel random-weighted loss function that assigns learning weights under Dirichlet distribution to prevent task dominance, which is able to speed up the convergence of the MTDL-based model and improve joint learning performance for automated diagnosis and severity assessment of COVID-19. Moreover, other optimization approaches that have shown superior performance in the field of computer vision, such as gradient normalization [234], dynamic weight averaging [235], and dynamic task prioritization [236], could be extended for future MTDL-based medical image computing and analysis.

4.4. Clinical application requirements and interpretability

The primary goal of developing MTDL-based medical image computing and analysis frameworks is to facilitate clinical decision-making, which needs objective and visualized evidence. Given that the artificial neural networks are somehow black boxes, the corresponding qualitative results (e.g., Dice coefficient) or quantitative results (e.g., segmentation mask) generated by the deep learning model may not be always reliable, hindering its application in clinical workflows. Thus, it is vital to enhance the interpretability of CNN-based models.

To better visualize the complex features learned by CNN-based models, some recent works combining CNN and Attention Mechanism (AM) have achieved outstanding results. AM can not only enhance the interpretability of CNN features but also improve the models' performance. In histopathological images processing [237–242], Huang et al. [240] have proposed an end-to-end ViT-AMCNet that incorporates an

attention mechanism-integrated convolution (AMC) block and a vision transformer (ViT) block to respectively produce a Gradient-weighted Class Activation Mapping (Grad-CAM) and a Rollout, which can visualize the feature knowledge learned by the AMC block and ViT-AMC block, respectively. Experimental results in Ref. [240] demonstrated that ViT-AMCNet significantly outperformed state-of-the-art methods. Importantly, the visualized interpretive maps are closer to the region of interest of concern by pathologists. Sun et al. [241] have harnessed CAM and guided backpropagation maps to visualize pixel-level morphological representations in histopathological images of the endometrium. Grad-CAM is also used to visualize the attention-guided discriminative regions (lymph nodes) where the MTDL model focuses [138]. Lu et al. [187] have computed the attention scores and visually interpreted the weights of different regions in the whole-slide image regarding the classification results of the MTDL model. In this way, clinic experts can examine the relatively important regions with high priority. Yan et al. [189] have proposed prior-aware MTDL for automated gland segmentation and tumor grading, where tissue prior information has been regarded as spatial attention in pathological interpretation. Moreover, except for the visualization, the CAM for one task can be fed into the other task subnet to guide its learning [35]. For brain structural MRI image processing, Lian et al. [243] have proposed a multi-task weakly-supervised attention network (MWAN) to jointly predict multiple clinical scores, and the visualization results show that the attention maps under MWAN are relatively more precise than those with conventional CAM. In the future, the MTDL model performed on medical images can take advantage of the approaches summarized in Ref. [244] to improve its interpretability. Besides, embedding the expert knowledge or clinical knowledge [245] into the MTDL model may make it more interpretable.

4.5. Data limitation and learning strategies beyond supervision

The unique traits of medical imaging data may influence the performance of MTDL-based medical image computing and analysis [2]. For example, the medical images are high-resolution, multi-modal, and difficult to collect because of patient privacy. The data acquisition and preprocessing procedures are inherited heterogeneous. The data distribution may be imbalanced. Some efforts in MTDL have been made for this challenging problem. For example, the focal loss utilized in Ref. [138] to alleviate positive and negative sample imbalance, the dice coefficient-based loss and modified weighted cross-entropy loss adopted in Ref. [50] to compensate for the imbalance between the foreground and background for the glioma segmentation, and the number of patients with IDH-wild gliomas and that with IDH-mutant gliomas. But a more effective counterbalancing mechanism is still desired to address this open issue.

The label information is limited, especially for MTDL models that may need more than two types of image annotation. Some studies have tried to obtain extra labeling information by constructing auxiliary labels based on expert annotations of the main task [98,214]. However, this approach can still not leverage the images without labels of the main task. Some MTDL models have introduced the technologies of weakly-supervised learning [32,246], semi-supervised learning [30,77,103,111], unsupervised learning [123], transfer learning [66], or contrastive learning [247] to alleviate the issue of limited labeling. Generally, weakly-supervised learning requires a small portion of strong labels, and the remaining majority of the data can be weakly labeled. Compared with the strong labels (e.g., lesion masks), weak labels (e.g., whether the lesion existent or not) are easier to annotate, significantly reducing the labeling costs. Semi-supervised learning can be applied when the labels are incomplete. The model is first pre-trained with the labeled data to generate pseudo or surrogate labels of unannotated images; Then, it is re-trained or fine-tuned by mixing up images with ground-truth and pseudo labels. Unsupervised learning does not rely on annotated images, which is popular in unsupervised domain adaptation [61,140,142]. Transfer learning seeks to improve the target domain tasks such as medical images with limited labels by taking advantage of knowledge learned from the source domain data, such as natural images with abundant annotations, which includes inductive [66] or unsupervised [248] transfer learning depending on the task similarities and data distributions between source and target domains. For instance, source and target domain data are the same diabetic retinopathy images [66]; In comparison, the source data is the public ImageNet, a large-scale natural image dataset with annotations, and the target data is non-stained grayscale sperm images [182]. Besides, multi-modality data fusion methods under the MTDL paradigm [38,200] can compensate for the limitation of data samples by taking full advantage of the collected multi-view data of each subject.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 81972160 and 81622025), the Startup Funds of Beijing Normal University, the China Scholarship Council (202106020141), and the Academic Excellence Foundation of BUAA for Ph.D. Students.

References

[1] H. Brody, Medical imaging, Nature 502 (7473) (2013). S81-S81.
[2] S.K. Zhou, et al., A review of deep learning in medical imaging: imaging traits, technology trends, case studies with progress highlights, and future promises, Proc. IEEE 109 (5) (2021) 820–838.
[3] D.S. Kermany, et al., Identifying medical diagnoses and treatable diseases by image-based deep learning, Cell 172 (5) (2018) 1122–1131, e9.
[4] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521 (7553) (2015) 436–444.
[5] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: surpassing human-level performance on ImageNet classification, in: IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1026–1034, 2015.
[6] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, DeepFace: closing the gap to human-level performance in face verification, in: IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1701–1708, 2014.
[7] A. Pucchio, E.A. Eisenhauer, F.Y. Moraes, Medical students need artificial intelligence and machine learning training, Nat. Biotechnol. 39 (3) (2021) 388–389.
[8] D. Shen, G. Wu, H.I. Suk, Deep learning in medical image analysis, Annu. Rev. Biomed. Eng. 19 (2017) 221–248 (in eng).
[9] O. Ozdemir, R.L. Russell, A.A. Berlin, A 3D probabilistic deep learning system for detection and diagnosis of lung cancer using low-dose CT scans, IEEE Trans. Med. Imag. 39 (5) (2020) 1419–1429.
[10] G. Luo, et al., Commensal correlation network between segmentation and direct area estimation for bi-ventricle quantification, Med. Image Anal. 59 (2020), 101591.
[11] Y. Xu, et al., Artificial intelligence: a powerful paradigm for scientific research, Innovation 2 (4) (2021), 100179.
[12] Y. Zhang, Q. Yang, A survey on multi-task learning, IEEE Trans. Knowl. Data Eng. 34 (12) (2022) 5586–5609.
[13] K.-H. Thung, C.-Y. Wee, A brief review on multi-task learning, Multimed. Tool. Appl. 77 (22) (2018) 29705–29725.
[14] Y. Zhang, Q. Yang, An overview of multi-task learning, Natl. Sci. Rev. 5 (1) (2018) 30–43.
[15] S. Vandenhende, S. Georgoulis, W.V. Gansbeke, M. Proesmans, D. Dai, L.V. Gool, Multi-task learning for dense prediction tasks: a survey, IEEE Trans. Pattern Anal. Mach. Intell. 44 (7) (2022) 3614–3633.
[16] S. Sosnin, M. Vashurina, M. Withnall, P. Karpov, M. Fedorov, I.V. Tetko, A survey of multi-task learning methods in chemoinformatics, tics 38 (4) (2019), 1800108.
[17] G. Litjens, et al., A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.
[18] H. Yu, L.T. Yang, Q. Zhang, D. Armstrong, M.J. Deen, Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives, Neurocomputing 444 (2021) 92–110.
[19] X. Xie, J. Niu, X. Liu, Z. Chen, S. Tang, S. Yu, A survey on incorporating domain knowledge into deep learning for medical image analysis, Med. Image Anal. 69 (2021), 101985.
[20] K. He, et al., HF-UNet: learning hierarchically inter-task relevance in multi-task U-net for accurate prostate segmentation in CT images, IEEE Trans. Med. Imag. 40 (8) (2021) 2118–2128.

[21] J. Zhao, W. Lv, B. Du, J. Ye, L. Sun, G. Xiong, Deep multi-task learning with flexible and compact architecture search, Int. J. Data Sci. Anal. (2021) 1–13.
[22] G. Wang, T. Song, Q. Dong, M. Cui, N. Huang, S. Zhang, Automatic ischemic stroke lesion segmentation from computed tomography perfusion images by image synthesis and attention-based deep neural networks, Med. Image Anal. 65 (2020), 101787.
[23] Z. Zhou, Y. Wang, Y. Guo, Y. Qi, J. Yu, Image quality improvement of hand-held ultrasound devices with a two-stage generative adversarial network, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 67 (1) (2020) 298–311.
[24] S. Tian, et al., DCPR-GAN: dental crown prosthesis restoration using two-stage generative adversarial networks, IEEE J. Biomed. Health Inform. 26 (1) (2021) 151–160.
[25] Y. Zhao, et al., Two-stream graph convolutional network for intra-oral scanner image segmentation, IEEE Trans. Med. Imag. 41 (4) (2021) 826–835.
[26] T.H. Wu, et al., Two-stage mesh deep learning for automated tooth segmentation and landmark localization on 3D intraoral scans, IEEE Trans. Med. Imag. 41 (11) (2022) 3158–3166.
[27] A. Malhotra, et al., Multi-task driven explainable diagnosis of COVID-19 using chest X-ray images, Pattern Recogn. (2021), 108243.
[28] Z. Xue, B. Xin, D. Wang, X. Wang, Radiomics-enhanced multi-task neural network for non-invasive glioma subtyping and segmentation, Radiom. Radiogenom. Neuro-oncol. (2020) 81–90.
[29] X. Xu, et al., Asymmetric multi-task attention network for prostate bed segmentation in computed tomography images, Med. Image Anal. (2021), 102116.
[30] X. Wang, et al., Towards multi-center glaucoma OCT image screening with semi-supervised joint structure and function multi-task learning, Med. Image Anal. 63 (2020), 101695.
[31] X. Wang, et al., A hybrid network for automatic hepatocellular carcinoma segmentation in H&E-stained whole slide images, Med. Image Anal. 68 (2021), 101914.
[32] C. Playout, R. Duval, F. Cheriet, A novel weakly supervised multitask architecture for retinal lesions segmentation on fundus images, IEEE Trans. Med. Imag. 38 (10) (2019) 2434–2444.
[33] X. Wang, et al., Joint learning of 3D lesion segmentation and classification for explainable COVID-19 diagnosis, IEEE Trans. Med. Imag. 40 (9) (2021) 2463–2476.
[34] W. Chen, Q. Wang, D. Yang, X. Zhang, C. Liu, Y. Li, End-to-End multi-task learning for lung nodule segmentation and diagnosis, in: 2020 25th International Conference on Pattern Recognition (ICPR), IEEE, 2021, pp. 6710–6717.
[35] Y. Xie, J. Zhang, Y. Xia, C. Shen, A mutual bootstrapping model for automated skin lesion segmentation and classification, IEEE Trans. Med. Imag. 39 (7) (2020) 2482–2493.
[36] Z. Tang, et al., Deep learning of imaging phenotype and genotype for predicting overall survival time of glioblastoma patients, IEEE Trans. Med. Imag. 39 (6) (2020) 2100–2109.
[37] J.L. Wang, H. Farooq, H. Zhuang, A.K. Ibrahim, Segmentation of intracranial hemorrhage using semi-supervised multi-task attention-based U-net, Appl. Sci. 10 (9) (2020) 3297, 2076-3417.
[38] H. Kuang, B.K. Menon, S.I.L. Sohn, W. Qiu, EIS-Net: segmenting early infarct and scoring ASPECTS simultaneously on non-contrast CT of patients with acute ischemic stroke, Med. Image Anal. 70 (2021), 101984.
[39] J. Zhao, et al., Functional network connectivity (FNC)-based generative adversarial network (GAN) and its applications in classification of mental disorders, J. Neurosci. Methods 341 (2020) 108756, 108756.
[40] L. Cao, et al., Multi-task neural networks for joint hippocampus segmentation and clinical score regression, Multimed. Tool. Appl. 77 (22) (2018) 29669–29686.
[41] N. Nandakumar, et al., A multi-task deep learning framework to localize the eloquent cortex in brain tumor patients using dynamic functional connectivity, in: Machine Learning in Clinical Neuroimaging and Radiogenomics in Neuro-Oncology, Springer International Publishing, Cham, 2020, pp. 34–44.
[42] M. Hamghalam, T. Wang, B. Lei, High tissue contrast image synthesis via multistage attention-GAN: application to segmenting brain MR scans, Neural Network. 132 (2020) 43–52.
[43] K. Zeng, H. Zheng, C. Cai, Y. Yang, K. Zhang, Z. Chen, Simultaneous single- and multi-contrast super-resolution for brain MRI images based on a convolutional neural network, Comput. Biol. Med. 99 (2018) 133–141.
[44] G. Wang, et al., Synthesize high-quality multi-contrast magnetic resonance imaging from multi-echo acquisition using multi-task deep generative model, IEEE Trans. Med. Imag. 39 (10) (2020) 3089–3099.
[45] X. Gao, F. Shi, D. Shen, M. Liu, Task-induced pyramid and attention GAN for multimodal brain image imputation and classification in alzheimer's disease, IEEE J. Biomed. Health Inform. 26 (1) (2022) 36–43.
[46] A. Elazab, , et al.G.P.- Gan, Brain tumor growth prediction using stacked 3D generative adversarial networks from longitudinal MR Images, Neural Network. 132 (2020) 321–332.
[47] X. Zhou, et al., Enhancing magnetic resonance imaging-driven Alzheimer's disease classification performance using generative adversarial learning, Alzheimer's Res. Ther. 13 (1) (2021) 60.
[48] Y. Luo, et al., Edge-preserving MRI image synthesis via adversarial network with iterative multi-scale fusion, Neurocomputing 452 (2021) 63–77.
[49] X. Yang, W.T. Tang, G. Tjio, S.Y. Yeo, Y. Su, Automatic detection of anatomical landmarks in brain MR scanning using multi-task deep neural networks, Neurocomputing 396 (2020) 514–521.
[50] J. Cheng, J. Liu, H. Kuang, J. Wang, A fully automated multimodal MRI-based multi-task learning for glioma segmentation and IDH genotyping, IEEE Trans. Med. Imag. 41 (6) (2022) 1520–1532.
[51] A. Myronenko, 3D MRI brain tumor segmentation using autoencoder regularization, in: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Springer International Publishing, Cham, 2019, pp. 311–320.
[52] H. Huang, et al., A deep multi-task learning framework for brain tumor segmentation, in: eng (Ed.), Front. Oncol. 11 (2021) 690244, 690244.
[53] T. Estienne, et al., U-ReSNet: ultimate coupling of registration and segmentation with deep nets, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019, Springer International Publishing, Cham, 2019, pp. 310–319.
[54] S. Chen, G. Bortsova, A. García-Uceda Juárez, G. van Tulder, M. de Bruijne, Multi-task Attention-Based Semi-supervised Learning for Medical Image Segmentation, Springer International Publishing, Cham, 2019, pp. 457–465.
[55] S. Spasov, L. Passamonti, A. Duggento, P. Liò, N. Toschi, A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease, in: eng (Ed.), Neuroimage 189 (2019) 276–287.
[56] M. Liu, J. Zhang, E. Adeli, D. Shen, Joint classification and regression via deep multi-task multi-channel learning for alzheimer's disease diagnosis, IEEE Trans. Biomed. Eng. 66 (5) (2019) 1195–1206.
[57] C. Lian, M. Liu, L. Wang, D. Shen, Multi-task weakly-supervised attention network for dementia status estimation with structural MRI, IEEE Transact. Neural Networks Learn. Syst. 33 (8) (2022) 4056–4068.
[58] M. Decuyper, S. Bonte, K. Deblaere, R. Van Holen, Automated MRI based pipeline for segmentation and prediction of grade, IDH mutation and 1p19q co-deletion in glioma, Comput. Med. Imag. Graph. 88 (2021), 101831.
[59] T. Zhou, H. Fu, G. Chen, J. Shen, L. Shao, Hi-net: hybrid-fusion network for multi-modal MR image synthesis, IEEE Trans. Med. Imag. 39 (9) (2020) 2772–2781.
[60] D. Wang, C. Wang, L. Masters, M. Barnett, Masked multi-task network for case-level intracranial hemorrhage classification in brain CT volumes, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020, Springer International Publishing, Cham, 2020, pp. 145–154.
[61] D. Tomar, M. Lortkipanidze, G. Vray, B. Bozorgtabar, J.P. Thiran, Self-Attentive spatial adaptive normalization for cross-modality domain adaptation, IEEE Trans. Med. Imag. 40 (10) (2021) 2926–2938.
[62] N. Nandakumar, et al., Automated eloquent cortex localization in brain tumor patients using multi-task graph neural networks, Med. Image Anal. 74 (2021), 102203.
[63] C. Zhou, C. Ding, X. Wang, Z. Lu, D. Tao, One-pass multi-task networks with cross-task guided attention for brain tumor segmentation, IEEE Trans. Image Process. 29 (2020) 4516–4529.
[64] M. Liu, et al., A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease, Neuroimage 208 (2020), 116459.
[65] L. Ju, et al., Synergic adversarial label learning for grading retinal diseases via knowledge distillation and multi-task learning, IEEE J. Biomed. Health Inform. 25 (10) (2021) 3709–3720.
[66] Y. Zhou, B. Wang, L. Huang, S. Cui, L. Shao, A benchmark for studying diabetic retinopathy: segmentation, grading, and transferability, IEEE Trans. Med. Imag. 40 (3) (2021) 818–828.
[67] Z. Wang, X. Jiang, J. Liu, K.T. Cheng, X. Yang, Multi-task siamese network for retinal artery/vein separation via deep convolution along vessel, IEEE Trans. Med. Imag. 39 (9) (2020) 2904–2919.
[68] J. Duan, et al., Automatic 3D bi-ventricular segmentation of cardiac images by a shape-refined multi-task deep learning approach, IEEE Trans. Med. Imag. 38 (9) (2019) 2151–2164.
[69] X. Wang, M. Xu, J. Zhang, L. Jiang, L. Li, Deep multi-task learning for diabetic retinopathy grading in fundus images, Proc. AAAI Conf. Artif. Intell. 35 (4) (2021) 2826–2834.
[70] X. Li, et al., Rotation-oriented collaborative self-supervised learning for retinal disease diagnosis, IEEE Trans. Med. Imag. 40 (9) (2021) 2284–2294.
[71] J. Wang, Y. Bai, B. Xia, Simultaneous diagnosis of severity and features of diabetic retinopathy in fundus photography using deep learning, IEEE J. Biomed. Health Inform. 24 (12) (2020) 3397–3407.
[72] S. Chelaramani, M. Gupta, V. Agarwal, P. Gupta, R. Habash, Multi-task knowledge distillation for eye disease prediction, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021, pp. 3983–3993.
[73] H. Xie, et al., AMD-GAN: attention encoder and multi-branch structure based generative adversarial networks for fundus disease detection from scanning laser ophthalmoscopy images, Neural Network. 132 (2020) 477–490.
[74] R. Zhang, et al., Biomarker localization by combining CNN classifier and generative adversarial network, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019, Springer International Publishing, Cham, 2019, pp. 209–217.
[75] B. Murugesan, K. Sarveswaran, S.M. Shankaranarayana, K. Ram, J. Joseph, M. Sivaprakasam, Psi-Net: shape and boundary aware joint multi-task deep network for medical image segmentation, in: 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2019, pp. 7223–7226, 2019.
[76] L. Yang, H. Wang, Q. Zeng, Y. Liu, G. Bian, A hybrid deep segmentation network for fundus vessels via deep-learning framework, Neurocomputing 448 (2021) 168–178.
[77] A. Foo, W. Hsu, M.L. Lee, G. Lim, T.Y. Wong, Multi-task learning for diabetic retinopathy grading and lesion segmentation, Proc. AAAI Conf. Artif. Intell. 34 (2020) 13267–13272, 08.

[78] Y. Zhou, X. He, S. Cui, F. Zhu, L. Liu, L. Shao, High-resolution diabetic retinopathy image synthesis manipulated by grading and lesions, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019, Springer International Publishing, Cham, 2019, pp. 505–513.
[79] X. Li, X. Hu, L. Yu, L. Zhu, C.-W. Fu, P.-A. Heng, CANet: cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading, IEEE Trans. Med. Imag. 39 (5) (2019) 1483–1493.
[80] W. Ma, S. Yu, K. Ma, J. Wang, X. Ding, Y. Zheng, Multi-task Neural Networks with Spatial Activation for Retinal Vessel Segmentation and Artery/Vein Classification, Springer International Publishing, Cham, 2019, pp. 769–778.
[81] T. He, J. Hu, Y. Song, J. Guo, Z. Yi, Multi-task learning for the segmentation of organs at risk with label dependence, Med. Image Anal. 61 (12) (2020), 101666.
[82] Y.H. Wu, et al., JCS: an explainable COVID-19 diagnosis system by joint classification and segmentation, IEEE Trans. Image Process. 30 (2021) 3113–3126.
[83] W. Liu, X. Liu, H. Li, M. Li, X. Zhao, Z. Zhu, Integrating lung parenchyma segmentation and nodule detection with deep multi-task learning, IEEE J. Biomed. Health Inform. 25 (8) (2021) 3073–3081.
[84] S. Marques, F. Schiavo, C.A. Ferreira, J. Pedrosa, A. Cunha, A. Campilho, A multi-task CNN approach for lung nodule malignancy classification and characterization, Expert Syst. Appl. 184 (2021), 115469.
[85] B. Lou, et al., An image-based deep learning framework for individualising radiotherapy dose: a retrospective analysis of outcome prediction, The Lancet Digit. Health 1 (3) (2019) e136–e147.
[86] M. Goncharov, et al., CT-Based COVID-19 triage: deep multitask learning improves joint identification and severity quantification, Med. Image Anal. 71 (2021), 102054.
[87] S. Park, W. Jeong, Y.S. Moon, X-Ray image segmentation using multi-task learning, KSII Trans. Internet and Inform. Syst. (TIIS) 14 (3) (2020) 1104–1120.
[88] P. Zhai, Y. Tao, H. Chen, T. Cai, J. Li, Multi-task learning for lung nodule classification on chest CT, IEEE Access 8 (2020) 180317–180327.
[89] X. Huang, W. Sun, T.-L. Tseng, C. Li, W. Qian, Fast and fully-automated detection and segmentation of pulmonary nodules in thoracic CT scans using deep convolutional neural networks, Comput. Med. Imag. Graph. 74 (2019) 25–36.
[90] D. Yu, et al., Detection of peripherally inserted central catheter (PICC) in chest X-ray images: a multi-task deep learning model, Comput. Methods Progr. Biomed. 197 (2020), 105674.
[91] J. Lian, et al., A structure-aware relation network for thoracic diseases detection and segmentation, in: eng (Ed.), IEEE Trans. Med. Imag. 40 (8) (2021) 2042–2052.
[92] A. Amyar, R. Modzelewski, H. Li, S. Ruan, Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: classification and segmentation, Comput. Biol. Med. 126 (2020), 104037.
[93] S. Park, et al., Multi-task vision transformer using low-level chest X-ray feature corpus for COVID-19 diagnosis and severity quantification, Med. Image Anal. 75 (2022), 102299.
[94] K. He, et al., Synergistic learning of lung lobe segmentation and hierarchical multi-instance classification for automated severity assessment of COVID-19 in CT images, Pattern Recogn. 113 (2021), 107828.
[95] A. Ghazipour, B. Veasey, A. Seow, A.A. Amini, Joint learning for deformable registration and malignancy classification of lung nodules, in: 2021 IEEE 18th International Symposium on Biomedical Imaging, ISBI), 2021, pp. 1807–1811.
[96] S. Mori, R. Hirai, Y. Sakata, Simulated four-dimensional CT for markerless tumor tracking using a deep learning network with multi-task learning, Phys. Med. 80 (2020) 151–158.
[97] Y. Yu, et al., Determining the invasiveness of ground-glass nodules using a 3D multi-task network, Eur. Radiol. 31 (9) (2021) 7162–7171.
[98] J. Liu, et al., RPLS-Net: pulmonary lobe segmentation based on 3D fully convolutional networks and multi-task learning, Int. J. Comput. Assist. Radiol. Surg. 16 (6) (2021) 895–904.
[99] N. Khosravan, H. Celik, B. Turkbey, E.C. Jones, B. Wood, U. Bagci, A collaborative computer aided diagnosis (C-CAD) system with eye-tracking, sparse attentional model, and deep learning, Med. Image Anal. 51 (2019) 101–115.
[100] L. Chenyang, S.C. Chan, A joint detection and recognition approach to lung cancer diagnosis from CT images with label uncertainty, IEEE Access 8 (2020) 228905–228921.
[101] Z. Tan, J. Feng, J. Zhou, Multi-task learning network for landmark detection in anatomical tree structures, in: 2021 IEEE 18th International Symposium on Biomedical Imaging, ISBI), 2021, pp. 1975–1979.
[102] L. Liu, Q. Dou, H. Chen, J. Qin, P.A. Heng, Multi-task deep model with margin ranking loss for lung nodule analysis, IEEE Trans. Med. Imag. 39 (3) (2020) 718–728.
[103] A.-A.-Z. Imran, D. Terzopoulos, Semi-supervised multi-task learning with chest X-ray images, in: Machine Learning in Medical Imaging, Springer International Publishing, Cham, 2019, pp. 151–159.
[104] B. Wu, Z. Zhou, J. Wang, Y. Wang, Joint learning for pulmonary nodule segmentation, attributes and malignancy prediction, in: IEEE 15th International Symposium on Biomedical Imaging, ISBI 2018), 2018, pp. 1109–1113, 2018.
[105] H. Tang, C. Zhang, X. Xie, NoduleNet: decoupled false positive reduction for pulmonary nodule detection and segmentation, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019, Springer International Publishing, Cham, 2019, pp. 266–274.
[106] L. Zhang, et al., Knowledge-guided multi-task attention network for survival risk prediction using multi-center computed tomography images, Neural Network. 152 (2022) 394–406.
[107] A. Amyar, R. Modzelewski, P. Vera, V. Morard, S. Ruan, Multi-task multi-scale learning for outcome prediction in 3D PET images, Comput. Biol. Med. 151 (2022), 106208.
[108] C. Xu, J. Howey, P. Ohorodnyk, M. Roth, H. Zhang, S. Li, Segmentation and quantification of infarction without contrast agents via spatiotemporal generative adversarial learning, Med. Image Anal. 59 (2020), 101568.
[109] K. Ta, S.S. Ahn, J.C. Stendahl, A.J. Sinusas, J.S. Duncan, A semi-supervised joint network for simultaneous left ventricular motion tracking and segmentation in 4D echocardiography, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2020, pp. 468–477.
[110] J. Chen, P. Zhang, H. Liu, L. Xu, H. Zhang, Spatio-temporal multi-task network cascade for accurate assessment of cardiac CT perfusion, Med. Image Anal. 74 (2021), 102207.
[111] S. Li, C. Zhang, X. He, Shape-aware semi-supervised 3d semantic segmentation for medical images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2020, pp. 552–561.
[112] Z. Guo, et al., DeepCenterline: a multi-task fully convolutional network for centerline extraction, in: Information Processing in Medical Imaging, Springer International Publishing, Cham, 2019, pp. 441–453.
[113] S. Vesal, M. Gu, A. Maier, N. Ravikumar, Spatio-temporal multi-task learning for cardiac MRI left ventricle quantification, IEEE J. Biomed. Health Inform. 25 (7) (2021) 2698–2709.
[114] W. Zhang, et al., Multi-task learning with multi-view weighted fusion attention for artery-specific calcification analysis, Inf. Fusion 71 (2021) 64–76.
[115] X. Du, R. Tang, S. Yin, Y. Zhang, S. Li, Direct segmentation-based full quantification for left ventricle via deep multi-task regression learning network, IEEE J. Biomed. Health Inform. 23 (3) (2019) 942–948.
[116] R. Ge, et al., K-net: integrate left ventricle segmentation and direct quantification of paired echo sequence, IEEE Trans. Med. Imag. 39 (5) (2020) 1690–1702.
[117] W. Xue, G. Brahm, S. Pandey, S. Leung, S. Li, Full left ventricle quantification via deep multitask relationships learning, Med. Image Anal. 43 (2018) 54–65.
[118] J. Chen, et al., JAS-GAN: generative adversarial network based joint atrium and scar segmentations on unbalanced atrial targets, IEEE J. Biomed. Health Inform. 26 (1) (2022) 103–114.
[119] W. Wang, Y. Wang, Y. Wu, T. Lin, S. Li, B. Chen, Quantification of full left ventricular metrics via deep regression learning with contour-guidance, IEEE Access 7 (2019) 47918–47928.
[120] Y. He, et al., Deep complementary joint model for complex scene registration and few-shot segmentation on medical images, in: Computer Vision – ECCV 2020, Springer International Publishing, Cham, 2020, pp. 770–786.
[121] X. Huang, Y. Tian, S. Zhao, T. Liu, W. Wang, Q. Wang, Direct full quantification of the left ventricle via multitask regression and classification, Appl. Intell. 51 (8) (2021) 5745–5758.
[122] K. Wang, B. Zhan, Y. Luo, J. Zhou, X. Wu, Y. Wang, Multi-task curriculum learning for semi-supervised medical image segmentation, in: 2021 IEEE 18th International Symposium on Biomedical Imaging, ISBI), 2021, pp. 925–928.
[123] C. Qin, et al., Joint learning of motion estimation and segmentation for cardiac MR image sequences, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018, Springer International Publishing, Cham, 2018, pp. 472–480.
[124] C. Yu, et al., Multi-level Multi-type Self-Generated Knowledge Fusion for Cardiac Ultrasound Segmentation, Information Fusion, 2022.
[125] C. Chen, W. Bai, D. Rueckert, Multi-task learning for left atrial segmentation on GE-MRI, in: Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges, Springer International Publishing, Cham, 2019, pp. 292–301.
[126] G.A. Bello, et al., Deep-learning cardiac motion analysis for human survival prediction, Nat. Mach. Intell. 1 (2) (2019) 95–104.
[127] M.H. Jafari, et al., Automatic biplane left ventricular ejection fraction estimation with mobile point-of-care ultrasound using multi-task learning and adversarial training, Int. J. Comput. Assist. Radiol. Surg. 14 (6) (2019) 1027–1037.
[128] M. Zreik, R.W.v. Hamersvelt, J.M. Wolterink, T. Leiner, M.A. Viergever, I. Išgum, A recurrent CNN for automatic detection and classification of coronary artery plaque and stenosis in coronary CT angiography, IEEE Trans. Med. Imag. 38 (7) (2019) 1588–1598.
[129] R. Chen, C. Xu, Z. Dong, Y. Liu, X. Du, DeepCQ: deep multi-task conditional quantification network for estimation of left ventricle parameters, Comput. Methods Progr. Biomed. 184 (2020), 105288.
[130] C. Yu, et al., Multitask learning for estimating multitype cardiac indices in MRI and CT based on adversarial reverse mapping, IEEE Transact. Neural Networks Learn. Syst. 32 (2) (2020) 493–506.
[131] Z. Zhu, et al., Lymph node gross tumor volume detection and segmentation via distance-based gating using 3D CT/PET imaging in radiotherapy, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020, Springer International Publishing, Cham, 2020, pp. 753–762.
[132] A. Grimwood, H. McNair, Y. Hu, E. Bonmati, D. Barratt, E.J. Harris, Assisted probe positioning for ultrasound guided radiotherapy using image sequence classification, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020, Springer International Publishing, Cham, 2020, pp. 544–552.
[133] S. Ramesh, et al., Multi-task temporal convolutional networks for joint recognition of surgical phases and steps in gastric bypass procedures, Int. J. Comput. Assist. Radiol. Surg. 16 (7) (2021) 1111–1119.
[134] B. Du, J. Liao, B. Turkbey, P. Yan, Multi-task learning for registering images with large deformation, IEEE J. Biomed. Health Inform. 25 (5) (2021) 1624–1633.
[135] Z. Xu, et al., Less is more: simultaneous view classification and landmark detection for abdominal ultrasound images, in: Medical Image Computing and

Computer Assisted Intervention – MICCAI 2018, Springer International Publishing, Cham, 2018, pp. 711–719.
[136] Q.-P. Liu, X. Xu, F.-P. Zhu, Y.-D. Zhang, X.-S. Liu, Prediction of prognostic risk factors in hepatocellular carcinoma with transarterial chemoembolization using multi-modal multi-task deep learning, EClinicalMedicine 23 (2020), 100379.
[137] J. Yao, et al., DeepPrognosis: preoperative prediction of pancreatic cancer survival and surgical margin via comprehensive understanding of dynamic contrast-enhanced CT imaging and tumor-vascular contact parsing, Med. Image Anal. 73 (2021), 102150.
[138] Y. Zhang, et al., 3D multi-attention guided multi-task learning network for automatic gastric tumor segmentation and lymph node classification, IEEE Trans. Med. Imag. 40 (6) (2021) 1618–1631.
[139] Y. Huo, et al., SynSeg-net: synthetic segmentation without target modality ground truth, IEEE Trans. Med. Imag. 38 (4) (2019) 1016–1025.
[140] W. Zeng, et al., Accurate 3d kidney segmentation using unsupervised domain translation and adversarial networks, in: 2021 IEEE 18th International Symposium on Biomedical Imaging, ISBI), 2021, pp. 598–602.
[141] B. Zhou, Z. Augenfeld, J. Chapiro, S.K. Zhou, C. Liu, J.S. Duncan, Anatomy-guided multimodal registration by learning segmentation without ground truth: application to intraprocedural CBCT/MR liver segmentation and registration, in: Medical Image Analysis vol. 71, 2021.
[142] J. Jiang, et al., PSIGAN: joint probabilistic segmentation and image distribution matching for unpaired cross-modality adaptation-based MRI segmentation, IEEE Trans. Med. Imag. 39 (12) (2020) 4071–4084.
[143] M. Zhu, Z. Chen, Y. Yuan, DSI-net: deep synergistic interaction network for joint classification and segmentation with endoscope images, in: eng (Ed.), IEEE Trans. Med. Imag. 40 (12) (2021) 3315–3325.
[144] M. Fang, et al., *Using Multi-Task Learning to Improve Diagnostic Performance of Convolutional Neural Networks* (SPIE Medical Imaging), SPIE, 2019.
[145] D. Wei, et al., SLIR: synthesis, localization, inpainting, and registration for image-guided thermal ablation of liver tumors, Med. Image Anal. 65 (2020), 101763.
[146] F. Liu, et al., JSSR: a joint synthesis, segmentation, and registration system for 3D multi-modal image alignment of large-scale pathological CT scans, in: Computer Vision – ECCV 2020, Springer International Publishing, Cham, 2020, pp. 257–274.
[147] S. Wang, K. He, D. Nie, S. Zhou, Y. Gao, D. Shen, CT male pelvic organ segmentation using fully convolutional networks with boundary sensitive representation, Med. Image Anal. 54 (2019) 168–178.
[148] Z. Feng, D. Nie, L. Wang, D. Shen, Semi-supervised learning for pelvic MR image segmentation based on multi-task residual fully convolutional networks, in: IEEE 15th International Symposium on Biomedical Imaging, ISBI 2018), 2018, pp. 885–888, 2018.
[149] Y. Ruan, et al., MB-FSGAN: joint segmentation and quantification of kidney tumor on CT by the multi-branch feature sharing generative adversarial network, Med. Image Anal. 64 (2020), 101721.
[150] Q. Wang, W. Xue, X. Zhang, F. Jin, J. Hahn, Pixel-wise body composition prediction with a multi-task conditional generative adversarial network, J. Biomed. Inf. 120 (2021), 103866.
[151] D. Keshwani, Y. Kitamura, S. Ihara, S. Iizuka, E. Simo-Serra, TopNet: topology preserving metric learning for vessel tree reconstruction and labelling, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020, Springer International Publishing, Cham, 2020, pp. 14–23.
[152] D. Keshwani, Y. Kitamura, Y. Li, Computation of total kidney volume from CT images in autosomal dominant polycystic kidney disease using multi-task 3D convolutional neural networks, in: Machine Learning in Medical Imaging, Springer International Publishing, Cham, 2018, pp. 380–388.
[153] J. Xue, et al., Cascaded MultiTask 3-D fully convolutional networks for pancreas segmentation, in: eng (Ed.), IEEE Trans. Cybern. 51 (4) (2021) 2153–2165.
[154] C. Jin, et al., Predicting treatment response from longitudinal images using multi-task deep learning, Nat. Commun. 12 (1) (2021) 1851.
[155] M.S. Elmahdy, et al., Joint registration and segmentation via multi-task learning for adaptive radiotherapy of prostate cancer, IEEE Access 9 (2021) 95551–95568.
[156] J. Yao, Y. Shi, L. Lu, J. Xiao, L. Zhang, DeepPrognosis: preoperative prediction of pancreatic cancer survival and surgical margin via contrast-enhanced CT imaging, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020, Springer International Publishing, Cham, 2020, pp. 272–282.
[157] F. Kordon, et al., Multi-task localization and segmentation for X-ray guided planning in knee surgery, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019, Springer International Publishing, Cham, 2019, pp. 622–630.
[158] X. Hu, et al., Joint landmark and structure learning for automatic evaluation of developmental dysplasia of the hip, IEEE J. Biomed. Health Inform. 26 (1) (2021) 345–358.
[159] M.A. Kaloi, X. Wang, K. He, Multi-task deep learning for child gender and age determination on hand radiographs, in: Biometric Recognition, Springer International Publishing, Cham, 2019, pp. 396–404.
[160] D. Zhang, J. Wang, J.H. Noble, B.M. Dawant, HeadLocNet: deep convolutional neural networks for accurate classification and multi-landmark localization of head CTs, Med. Image Anal. 61 (2020), 101659.
[161] D. Zhang, B. Chen, S. Li, Sequential conditional reinforcement learning for simultaneous vertebral body detection and segmentation with modeling the spine anatomy, Med. Image Anal. 67 (2021), 101861.
[162] A.A.Z. Imran, et al., Partly supervised multi-task learning, in: 2020 19th IEEE International Conference on Machine Learning and Applications, ICMLA), 2020, pp. 769–774.
[163] Z. Huang, et al., DA-GAN: learning structured noise removal in ultrasound volume projection imaging for enhanced spine segmentation, in: 2021 IEEE 18th International Symposium on Biomedical Imaging, ISBI), 2021, pp. 770–774.
[164] C. Tan, L. Zhao, Z. Yan, K. Li, D. Metaxas, Y. Zhan, Deep multi-task and task-specific feature learning network for robust shape preserved organ segmentation, in: IEEE 15th International Symposium on Biomedical Imaging, ISBI 2018), 2018, pp. 1221–1224, 2018.
[165] P. Wang, M. Vives, V.M. Patel, I. Hacihaliloglu, Robust bone shadow segmentation from 2D ultrasound through task decomposition, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020, Springer International Publishing, Cham, 2020, pp. 805–814.
[166] C.E. von Schacky, et al., Development and validation of a multitask deep learning model for severity grading of hip osteoarthritis features on radiographs, Radiology 295 (1) (2020) 136–145 (in eng).
[167] S. Sukegawa, et al., Multi-task deep learning model for classification of dental implant brand and treatment stage using dental panoramic radiograph images, Biomolecules 11 (6) (2021).
[168] X. Fu, G. Yang, K. Zhang, N. Xu, J. Wu, An automated estimator for Cobb angle measurement using multi-task networks, Neural Comput. Appl. 33 (10) (2021) 4755–4761.
[169] Y. Gao, C. Liu, L. Zhao, Multi-resolution path CNN with deep supervision for intervertebral disc localization and segmentation, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019, Springer International Publishing, Cham, 2019, pp. 309–317.
[170] Y. Hong, B. Wei, Z. Han, X. Li, Y. Zheng, S. Li, MMCL-Net: spinal disease diagnosis in global mode using progressive multi-task joint learning, Neurocomputing 399 (2020) 307–316.
[171] R. Zhang, X. Xiao, Z. Liu, Y. Li, S. Li, MRLN: multi-task relational learning network for MRI vertebral localization, identification, and segmentation, IEEE J. Biomed. Health Inform. 24 (10) (2020) 2902–2911.
[172] F. Liu, Y. Jonmohamadi, G. Maicas, A.K. Pandey, G. Carneiro, Self-supervised depth estimation to regularise semantic segmentation in knee arthroscopy, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2020, pp. 594–603.
[173] N. Bayramoglu, J. Kannala, J. Heikkilä, Deep learning for magnification independent breast cancer histopathology image classification, in: 23rd International Conference on Pattern Recognition, ICPR), 2016, pp. 2440–2445, 2016.
[174] H. Chen, X. Qi, L. Yu, P.-A. Heng, DCAN: deep contour-aware networks for accurate gland segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2487–2496.
[175] S. Mehta, E. Mercan, J. Bartlett, D. Weaver, J.G. Elmore, L. Shapiro, Y-Net, Joint segmentation and classification for diagnosis of breast biopsy images, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018, Springer International Publishing, Cham, 2018, pp. 893–901.
[176] S. Takahama, et al., Multi-stage pathological image classification using semantic segmentation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 10702–10711.
[177] C.F. Koyuncu, G.N. Gunesli, R. Cetin-Atalay, C. Gunduz-Demir, DeepDistance: a multi-task deep regression model for cell detection in inverted microscopy images, Med. Image Anal. 63 (2020), 101720.
[178] H. Chen, X. Qi, L. Yu, Q. Dou, J. Qin, P.-A. Heng, DCAN: deep contour-aware networks for object instance segmentation from histology images, Med. Image Anal. 36 (2017) 135–146.
[179] H. Yu, et al., Large-scale gastric cancer screening and localization using multi-task deep neural network, Neurocomputing 448 (2021) 290–300.
[180] T. Peng, M. Boxberg, W. Weichert, N. Navab, C. Marr, Multi-task Learning of a Deep K-Nearest Neighbour Network for Histopathological Image Classification and Retrieval, Springer International Publishing, Cham, 2019, pp. 676–684.
[181] L. Li, et al., Multi-task deep learning for fine-grained classification and grading in breast cancer histopathological images, Multimed. Tool. Appl. 79 (21) (2020) 14509–14528.
[182] A. Abbasi, E. Miahi, S.A. Mirroshandel, Effect of deep transfer and multi-task learning on sperm abnormality detection, Comput. Biol. Med. 128 (2021), 104121.
[183] Y. Zhu, E. Meijering, Automatic improvement of deep learning-based cell segmentation in time-lapse microscopy by neural architecture search, Bioinformatics 37 (24) (2021) 4844–4850.
[184] S. Graham et al., "One model is all you need: multi-task learning enables simultaneous histology image segmentation and classification," Med. Image Anal., vol. 83, p. 102685, 2023.
[185] Z. Gao et al., "A semi-supervised multi-task learning framework for cancer classification with weak annotation in whole-slide images," Med. Image Anal., vol. 83, p. 102652, 2023.
[186] M. Dabass, S. Vashisth, R. Vig, MTU: a multi-tasking U-net with hybrid convolutional learning and attention modules for cancer classification and gland Segmentation in Colon Histopathological Images, Comput. Biol. Med. 150 (2022), 106095.
[187] M.Y. Lu, et al., AI-based pathology predicts origins for cancers of unknown primary, Nature 594 (7861) (2021) 106–110.
[188] J. Zhang, et al., Chromosome classification and straightening based on an interleaved and multi-task network, IEEE J. Biomed. Health Inform. 25 (8) (2021) 3240–3251.
[189] C. Yan, J. Xu, J. Xie, C. Cai, H. Lu, Prior-aware CNN with multi-task learning for colon images analysis, in: 2020 IEEE 17th International Symposium on Biomedical Imaging, ISBI), 2020.

[190] M. Fan, T. Chakraborti, E.I.C. Chang, Y. Xu, J. Rittscher, Microscopic fine-grained instance classification through deep attention, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020, Springer International Publishing, Cham, 2020, pp. 490–499.
[191] J. Qin, Y. He, J. Ge, Y. Liang, A multi-task feature fusion model for cervical cell classification, IEEE J. Biomed. Health Inform. 26 (9) (2022) 4668–4678.
[192] M. Machoy, J. Seeliger, L. Szyszka-Sommerfeld, R. Koprowski, T. Gedrange, K. Woźniak, The use of optical coherence tomography in dental diagnostics: a state-of-the-art review, Journal of healthcare engineering 2017 (2017).
[193] X. Zhou, Y. Gan, J. Xiong, D. Zhang, Q. Zhao, Z. Xia, A method for tooth model reconstruction based on integration of multimodal images, J. Healthcare Eng. 2018 (2018).
[194] M. Rajee, C. Mythili, Gender classification on digital dental X-ray images using deep convolutional neural network, Biomed. Signal Process Control 69 (2021), 102939.
[195] C.H. Wu, W.H. Tsai, Y.H. Chen, J.K. Liu, Y.N. Sun, Model-based orthodontic assessments for dental panoramic radiographs, IEEE J. Biomed. Health Inform. 22 (2) (2018) 545–551.
[196] S. Tian, et al., Efficient computer-aided design of dental inlay restoration: a deep adversarial framework, IEEE Trans. Med. Imag. 40 (9) (2021) 2415–2427.
[197] Y. Lai, et al., Lcanet: learnable connected attention network for human identification using dental images, IEEE Trans. Med. Imag. 40 (3) (2020) 905–915.
[198] M. Chung, et al., Automatic registration between dental cone-beam CT and scanned surface via deep pose regression neural networks and clustered similarities, IEEE Trans. Med. Imag. 39 (12) (2020) 3900–3909.
[199] J. Zhang, et al., Joint Craniomaxillofacial Bone Segmentation and Landmark Digitization by Context-Guided Fully Convolutional Networks, Springer International Publishing, Cham, 2017, pp. 720–728.
[200] J. Kawahara, S. Daneshvar, G. Argenziano, G. Hamarneh, Seven-point checklist and skin lesion classification using multitask multimodal neural nets, IEEE J. Biomed. Health Inform. 23 (2) (2019) 538–546.
[201] Z. Sobhaninia, et al., Fetal ultrasound image segmentation for measuring biometric parameters using multi-task deep learning, in: 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC), 2019, pp. 6545–6548, 2019.
[202] L. Liao, et al., Joint image quality assessment and brain extraction of fetal MRI using deep learning, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020, Springer International Publishing, Cham, 2020, pp. 415–424.
[203] Q. Xu, et al., Multi-task joint learning model for segmenting and classifying tongue images using a deep neural network, IEEE J. Biomed. Health Inform. 24 (9) (2020) 2481–2489.
[204] A.I.L. Namburete, W. Xie, M. Yaqub, A. Zisserman, J.A. Noble, Fully-automated alignment of 3D fetal brain ultrasound to a canonical reference space using multi-task learning, Med. Image Anal. 46 (2018) 1–14.
[205] L. Liu, J. Xu, Y. Huan, Z. Zou, S.C. Yeh, L.R. Zheng, A smart dental health-IoT Platform based on intelligent hardware, deep learning, and mobile terminal, IEEE J. Biomed. Health Inform. 24 (3) (2020) 898–906.
[206] Y. Chen, et al., Automatic segmentation of individual tooth in dental CBCT images from tooth surface map by a multi-task FCN, IEEE Access 8 (2020) 97296–97309.
[207] Z. Cui, et al., Hierarchical morphology-guided tooth instance segmentation from CBCT images, in: International Conference on Information Processing in Medical Imaging, Springer, 2021, pp. 150–162.
[208] Z. Cui, et al., TSegNet: an efficient and accurate tooth segmentation network on 3D dental model, Med. Image Anal. 69 (2021), 101949.
[209] Y. Lang, et al., Localization of Craniomaxillofacial landmarks on CBCT images using 3D mask R-CNN and local dependency learning, IEEE Trans. Med. Imag. 41 (10) (2022) 2856–2866.
[210] Z. Cui, et al., A fully automatic AI system for tooth and alveolar bone segmentation from cone-beam CT images, Nat. Commun. 13 (1) (2022) 1–11.
[211] C. Lian, et al., Multi-task dynamic transformer network for concurrent bone segmentation and large-scale landmark localization with dental CBCT, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2020, pp. 807–816.
[212] L. Song, J. Lin, Z.J. Wang, H. Wang, An end-to-end multi-task deep learning framework for skin lesion analysis, IEEE J. Biomed. Health Inform. 24 (10) (2020) 2912–2921.
[213] S. Chen, Z. Wang, J. Shi, B. Liu, N. Yu, A multi-task framework with feature passing module for skin lesion classification and segmentation, in: IEEE 15th International Symposium on Biomedical Imaging, ISBI 2018), 2018, pp. 1126–1129, 2018.
[214] L. Liu, Y.Y. Tsui, M. Mandal, Skin lesion segmentation using deep learning with auxiliary task, J. Imag. 7 (4) (2021) 67.
[215] M.V.S. de Cea, K. Diedrich, R. Bakalo, L. Ness, D. Richmond, Multi-task learning for detection and classification of cancer in screening mammography, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2020, pp. 241–250.
[216] X. Hou, Y. Bai, Y. Xie, Y. Li, Mass segmentation for whole mammograms via attentive multi-task learning framework, Phys. Med. Biol. 66 (10) (2021), 105015.
[217] R. Shen, K. Zhou, K. Yan, K. Tian, J. Zhang, Multicontext multitask learning networks for mass detection in mammogram, Med. Phys. 47 (4) (2020) 1566–1578.
[218] Y. Zhou, et al., Multi-task learning for segmentation and classification of tumors in 3D automated breast ultrasound images, Med. Image Anal. 70 (2021), 101918.
[219] F. Gao, H. Yoon, T. Wu, X. Chu, A feature transfer enabled multi-task deep learning model on medical imaging, Expert Syst. Appl. 143 (2020), 112957.
[220] G. Zhang, K. Zhao, Y. Hong, X. Qiu, K. Zhang, B. Wei, SHA-MTL: soft and hard attention multi-task learning for automated breast cancer ultrasound image segmentation and classification, Int. J. Comput. Assist. Radiol. Surg. 16 (10) (2021) 1719–1725.
[221] Z. Lin, et al., Multi-task learning for quality assessment of fetal head ultrasound images, Med. Image Anal. 58 (2019), 101548.
[222] S. Kyung, et al., Improved performance and robustness of multi-task representation learning with consistency loss between pretexts for intracranial hemorrhage identification in head CT, Med. Image Anal. 81 (2022), 102489.
[223] M. Meng, B. Gu, L. Bi, S. Song, D.D. Feng, J. Kim, DeepMTS: deep multi-task learning for survival prediction in patients with advanced nasopharyngeal carcinoma using pretreatment PET/CT, IEEE J. Biomed. Health Inform. 26 (9) (2022) 4497–4507.
[224] C.M. Tam, D. Zhang, B. Chen, T. Peters, S. Li, Holistic multitask regression network for multiapplication shape regression segmentation, Med. Image Anal. 65 (2020), 101783.
[225] Z. Huang, et al., Considering anatomical prior information for low-dose CT image enhancement using attribute-augmented Wasserstein generative adversarial networks, Neurocomputing 428 (2021) 104–115.
[226] A. Harouni, A. Karargyris, M. Negahdar, D. Beymer, T. Syeda-Mahmood, Universal multi-modal deep network for classification and segmentation of medical images, in: IEEE 15th International Symposium on Biomedical Imaging, ISBI 2018), 2018, pp. 872–876, 2018.
[227] S. Wang, et al., Global-Local attention network with multi-task uncertainty loss for abnormal lymph node detection in MR images, Med. Image Anal. 77 (2022), 102345.
[228] Y. Zhao, et al., Multi-view prediction of Alzheimer's disease progression with end-to-end integrated framework, J. Biomed. Inf. 125 (2022), 103978.
[229] Y. Gao, H. Bai, Z. Jie, J. Ma, K. Jia, W. Liu, Mtl-nas: task-agnostic neural architecture search towards general-purpose multi-task learning, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 11543–11552.
[230] C. Fifty, E. Amid, Z. Zhao, T. Yu, R. Anil, C. Finn, Efficiently Identifying Task Groupings for Multi-Task Learning, 2021 *arXiv preprint arXiv:2109.04617*.
[231] Y. Zou, et al., Multi-task deep learning based on T2-weighted images for predicting muscular-invasive bladder cancer, Comput. Biol. Med. 151 (2022), 106219.
[232] S. Liu, H. Wang, Y. Li, X. Li, G. Cao, W. Cao, AHU-MultiNet: adaptive loss balancing based on homoscedastic uncertainty in multi-task medical image segmentation network, Comput. Biol. Med. 150 (2022), 106157.
[233] G. Bao, et al., COVID-MTL: multitask learning with Shift3D and random-weighted loss for COVID-19 diagnosis and severity assessment, Pattern Recogn. 124 (2022), 108499.
[234] Z. Chen, V. Badrinarayanan, C.-Y. Lee, A. Rabinovich, Gradnorm: gradient normalization for adaptive loss balancing in deep multitask networks, in: International Conference on Machine Learning, PMLR, 2018, pp. 794–803.
[235] S. Liu, E. Johns, A.J. Davison, End-to-end multi-task learning with attention, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 1871–1880.
[236] M. Guo, A. Haque, D.-A. Huang, S. Yeung, L. Fei-Fei, Dynamic task prioritization for multitask learning, in: Proceedings of the European Conference on Computer Vision, ECCV), 2018, pp. 270–287.
[237] P. Huang, X. Tan, X. Zhou, S. Liu, F. Mercaldo, A. Santone, FABNet: fusion attention block and transfer learning for laryngeal cancer tumor grading in P63 IHC histopathology images, IEEE J. Biomed. Health Inform. 26 (4) (2021) 1696–1707.
[238] X. Zhou, C. Tang, P. Huang, S. Tian, F. Mercaldo, A. Santone, ASI-DBNet: an adaptive sparse interactive ResNet-vision transformer dual-branch network for the grading of brain cancer histopathological images, Interdiscipl. Sci. Comput. Life Sci. (2022) 1–17.
[239] X. Zhou, C. Tang, P. Huang, F. Mercaldo, A. Santone, Y. Shao, LPCANet: classification of laryngeal cancer histopathological images using a CNN with position attention and channel attention mechanisms, Interdiscipl. Sci. Comput. Life Sci. 13 (4) (2021) 666–682.
[240] P. Huang, et al., A ViT-AMC network with adaptive model fusion and multiobjective optimization for interpretable laryngeal tumor grading from histopathological images, IEEE Trans. Med. Imag. 42 (1) (2023) 15–28.
[241] H. Sun, X. Zeng, T. Xu, G. Peng, Y. Ma, Computer-aided diagnosis in histopathological images of the endometrium using a convolutional neural network and attention mechanisms, IEEE j. biomed. health inform. 24 (6) (2019) 1664–1676.
[242] P. Wang, et al., Cross-task extreme learning machine for breast cancer image classification with deep convolutional features, Biomed. Signal Process Control 57 (2020), 101789.
[243] C. Lian, M. Liu, L. Wang, D. Shen, Multi-task weakly-supervised attention network for dementia status estimation with structural mri, IEEE Transact. Neural Networks Learn. Syst. (2021) 1–13.
[244] Y. Zhang, P. Tiňo, A. Leonardis, K. Tang, A survey on neural network interpretability, IEEE Trans. Emerg. Topic. Comput. Int. 5 (5) (2021) 726–742.
[245] P. Deng, X. Han, X. Wei, L. Chang, Automatic classification of thyroid nodules in ultrasound images using a multi-task attention network guided by clinical knowledge, Comput. Biol. Med. 150 (2022), 106172.

[246] R. Ke, A. Bugeau, N. Papadakis, M. Kirkland, P. Schuetz, C.B. Schönlieb, Multi-task deep learning for image segmentation using recursive approximation tasks, IEEE Trans. Image Process. 30 (2021) 3555–3567.
[247] J. Li, et al., Multi-task contrastive learning for automatic CT and X-ray diagnosis of COVID-19, Pattern Recogn. 114 (2021), 107848.
[248] H. Chang, J. Han, C. Zhong, A.M. Snijders, J.H. Mao, Unsupervised transfer learning via multi-scale convolutional sparse coding for biomedical applications, IEEE Trans. Pattern Anal. Mach. Intell. 40 (5) (2018) 1182–1194.