International Journal of Computing Sciences Research (ISSN print: 2546-0552; ISSN online: 2546-115X)
Vol. 9, pp. 3418-3440
doi: 10.25147/ijcsr.2017.001.1.226
https://stepacademic.net

Short Paper

Multi-task Deep Learning in Medical Image Processing: A Systematic Review

Jackson Kamiri
Department of Information Technology, Murang'a University of Technology, Kenya
kamirijackson@gmail.com / jkamiri@mut.ac.ke
https://orcid.org/0000-0001-6899-1152
(Corresponding author)

Geoffrey M. Wambugu
Department of Information Technology, Murang'a University of Technology, Kenya
https://orcid.org/0000-0002-0250-0135

Aaron M. Oirere
Department of Computer Science, Murang'a University of Technology, Kenya
https://orcid.org/0000-0002-3167-9429

Date received: July 20, 2024
Date received in revised form: November 24, 2024
Date accepted: January 12, 2025

Recommended citation:

> Kamiri, J., Wambugu, G.M., & Oirere, A.M. (2025). Multi-task Deep Learning in Medical Image processing: A systematic review. *International Journal of Computing Sciences Research*, *9*, 3418-3440. https://doi.org/10.25147/ijcsr.2017.001.1.226

Abstract

Purpose — Multi-task learning (MTL) is a deep learning approach that aims to jointly learn two or more tasks with the goal of leveraging shared knowledge among the tasks. This study aimed to review existing MTL models in medical image processing to understand the current state of research, evaluate major breakthroughs, and analyze open gaps and future research direction.

Methodology — The study conducted a systematic literature review employing a search of peer-reviewed journal articles and conference proceedings. The articles were sourced from IEEE, ScienceDirect, PubMed, and Google Scholar databases. 52 primary papers published between 2016 and 2024 were considered in this study.

Results — The study's findings reveal that breakthroughs have been made in increasing the scope of task combinations in both homogenous and heterogenous tasks. Additionally, innovative architectural designs and learning methods have emerged. Although MTL has emerged as a panacea for medical image processing, some grey areas in research need to be addressed. They include task relatedness, scope of task combinations, generative MTL, and longitudinal MTL.

Conclusion — The study conducted a comprehensive analysis of multi-task models in medical image processing. The findings reveal breakthroughs in architecture, task combinations, and learning methods, and open gaps in this field. Metrics variability and proprietary datasets were the major limitations of this study.

Recommendations — Future researchers should focus on addressing the gaps identified in this study especially increasing the scope of MTL and designing more robust and highly generalizable neural networks for longitudinal MTL.

Research Implications — The review evaluates the current state of medical image processing using MTL, offering insights into both theoretical and practical aspects. These insights provide direction for future researchers to advance the field and for policymakers to support ethical data collection and sharing.

Keywords — multi-task learning, deep learning, medical image processing, computer vision, neural networks

INTRODUCTION

Multi-task learning (MTL) is a deep learning approach that aims at jointly learning two or more tasks together to leverage shared knowledge among the tasks. This shared knowledge helps to improve the performance of each task(Zhang & Yang, 2018). Traditionally deep learning models were trained to perform a single task; however, these traditional single-task models required a lot of training data which was not readily available in some domains such as medicine and bioinformatics. This gap led to the emergence of transfer learning where a deep neural network would be pre-trained using large-scale datasets such as ImageNet (Deng et al., 2009). The weights learned from ImageNet could then be transferred to performing new tasks that had smaller datasets such as medical image processing(Qummar et al., 2019). Some of the popular transfer learning networks that emerged and have been used in medical image processing include; VGG nets (Simonyan & Zisserman, 2015), Residual Networks (He et al., 2016), DenseNets (Huang et al., 2017), Alexnet (Krizhevsky et al., 2012), EfficientNet (M. Tan & Le, 2020), MobileNet (Howard et al., 2017) InceptionNet (Szegedy et al., 2015). The goal of transfer learning is to optimize a target task with the help of a source task (Zhang & Yang, 2022).

Although the concept of transfer learning enabled the development of deep learning solutions in domains where large datasets were not available, it faced challenges since only single-task models were developed from this. The negative transfer was also a major issue since the ImageNet dataset does not contain biomedical images. MTL seeks to address this gap by providing a way in which related tasks could be learned and optimized together. This mitigates overfitting by fostering a model's robustness and fostering the acquisition of a universal representation across multiple tasks (Ruder, 2017). Therefore, assuming that the tasks or a subset of them are related, then learning them together has been theoretically and empirically proven to improve performance as opposed to single-task learning (Zhang & Yang, 2018). Multi-task learning can be grouped into six categories namely; multi-task supervised learning, multi-task semi-supervised learning (Khosravan & Bagci, 2018), multi-task unsupervised learning, multi-task reinforcement learning, multi-task active learning, and multi-task online learning(Zhang & Yang, 2018).

Computer-aided medical image processing involves using computing technologies to process medical images. This has played a critical role in the diagnosis of various medical disorders from medical images (Rajpurkar et al., 2017; Yu et al., 2021), (Zhang et al., 2021). Medical images are unique in several ways, first unlike natural images, intensity, and variance are very crucial in medical images since they could mean the presence or absence of a biological feature (Taghanaki et al., 2021). Second, scale plays a critical role since knowing the pixel size of a feature in a medical image can help in computing the size of something like a tumor. There exist numerous medical imaging modalities that can be used to obtain medical images. They include X-ray, Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emitted Tomography (PET), Ultrasound, Microscopy, Retinal Imaging, Dermoscopy, Histopathology, X-ray mammography, Endoscopy, and X-ray Fluoroscopy, among others. Each of these modalities results in images with unique features. For instance, X-ray generates 2D images while MRI can generate 3D images. MRI has different modalities namely T1, T2, Flair, and T1ce(Çetiner & Metlek, 2023) which have varying levels of detail and complexity during processing.

The field of medical image processing can benefit greatly from multi-task learning. First, MTL can help learn an array of related tasks such as medical image segmentation, object detection, classification, and reconstruction. Joint learning helps leverage shared knowledge leading to improved performance and efficiency. Secondly, MTL can help deal with multi-modality in medical images through multi-modal fusion(Çetiner & Metlek, 2023). This involves combining different modalities to get a much-detailed understating of the organ under observation since each modality provides varying levels of detail, especially at the pixel level(Helland et al., 2023). Multi-task learning can be influenced by two factors which are task relatedness and task definition (Zhang & Yang, 2018; Ruder, 2017; Zhang & Yang, 2022). Therefore, there is a need for researchers to understand which types of MTL can be ideal for their study as well as how task combinations can be done concerning task relatedness and task definition. Failure to consider this can result in MTL models that are not generalizable, optimal, and efficient.

Although researchers such as (Buettner et al., 2020; Vetriselvi & Thenmozhi,2023; and Sohail et al., 2022) have done systematic reviews on the use of deep learning in medical imaging, to the best of our knowledge no researcher has done a comprehensive systematic review of multi-task learning in medical image processing. Therefore, considering the immense potential of MTL in medical imaging and the breakthroughs made in the field, there is a need for a study that provides a systematic review of MTL in medical image processing. This study was motivated to fill this gap in the literature by conducting a thorough systematic review of multi-task deep neural networks in medical image processing. A systematic literature review is an evidence-based way of evaluating and interpreting available research relevant to a particular research (Kitchenham et al., 2009). The study analyzes the latest developments and breakthroughs in this field thus providing a basis from which researchers can learn different cutting-edge MTL techniques in medical image processing.

METHODOLOGY

The researchers employed a Systematic literature review following the guidelines of (Kitchenham et al., 2007 as cited by Kitchenham et al., 2009). The guideline describes a three-phase process of conducting an evidence-based systematic literature review. The three phases are; planning the review, conducting the review, and reporting the findings of the review. The choice of this guideline was informed by the comprehensive nature of the guideline since according to (Kitchenham et al., 2009) the guideline combines three other guidelines, two books, and discussions with participants from different disciplines.

Research Questions

The study was guided by the following research questions (RQ)

1. RQ1: Which type of data and imaging modalities have the models used?
2. RQ2: What techniques have the existing MTL models used in medical image processing?
3. RQ3: Which task combinations have the models used?
4. RQ4: What are the outstanding gaps in the existing models?

The research questions were derived as part of the planning phase of the phase systematic literature review process. The four questions cover the critical aspects of MTL in medical image processing. The first question seeks to understand the types of data and imaging modalities that are used together in medical image processing. The second question seeks to understand how researchers formulate their MTL setup, the third question seeks to understand how different medical imaging tasks can be combined in an MTL setup and the fourth question explores the gaps that existing studies have not covered so that future work can address them.

Inclusion criteria

For a study to be included in this review, it needed to satisfy the following criteria: First, the paper must be using deep learning to do more than one task in medical image processing. Second, the study must be published in a peer-reviewed journal or conference proceedings. Third, the study must have a strong methodology and theoretical foundation. Fourth, the study must be a primary research paper. Fifth, the study must have reported its results using evaluation metrics.

Exclusion criteria

Any study that did not meet the inclusion criteria was excluded. In addition to that, studies older than eight years and studies not published in English were also not included. Where several versions of a paper excited the most recent version was considered.

Search process

To identify relevant papers, the researchers used a combination of searching for papers from research databases and identification of papers using references from included studies. The databases considered were Google Scholar, IEEE, PubMed, and Science Direct. The keywords used in the search were: "Multi-task learning in medical imaging", "Multi-task learning in medical image processing", "Multi-task deep learning in medical imaging", "Multi-task neural networks for medical imaging", "Multi-task neural networks for medical image processing", and "Deep learning in medical image processing". The choice of several keywords was informed by the need to reduce the risk of omission of relevant studies.

Quality Assessment

Kitchenham et al., (2009) argue that quality is subjective, therefore, the researchers in this study set the quality threshold as the ability of the study to answer all the research questions. To determine how well each study satisfied the quality criteria, the researchers used the following approach. To satisfy RQ1 the researchers read the methodology section documenting the data used. To satisfy RQ2 and RQ3 researchers read the Abstract, methodology, and results sections documenting the architecture, techniques used, and implementation details of the model. To satisfy RQ4 the researchers considered the methodology, results, discussion, future works of the study, and independent analysis.

Deviation from protocol

A protocol in a systematic literature review defines the methods that will be used in the review. It contains all elements of the review and some additional planning information (Kitchenham et al., 2009). The researchers in this study followed the defined methodology to ensure that the findings were not influenced by researcher biases.

RESULTS AND DISCUSSION

This section documents the findings of the study. It demonstrates how the research questions were answered in the study. A total of 247 articles were identified, 24 articles were removed as per the exclusion criteria because they were duplicates, 110 articles were also removed because they failed to meet the inclusion criteria, and 113 articles were screened through the quality assessment criteria out of which only 52 qualified to be included in this review. Figure 1 shows the article selection processes. The papers were published between 2016 and 2024. The data reveals that 36 of the papers (69.2%) included in the study were published between 2020 and 2021. Notably, this has been an active field since 2023 has four papers in the study. One of the papers in 2023 (Butoi et al., 2023) was identified as a ground-breaking study in medical image segmentation.

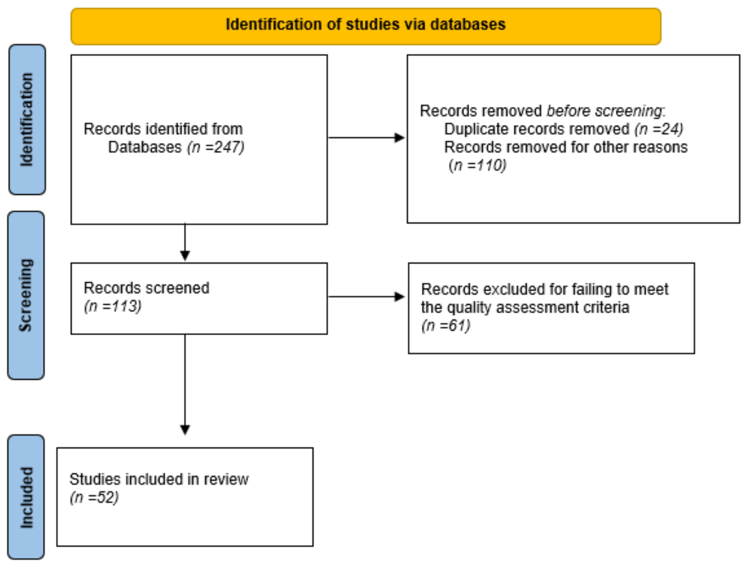

Figure 1: The selection process of articles that were included in the study

Data Extraction

The authors created a spreadsheet that was used to extract data. Each paper was placed in its individual worksheet and the following elements of the paper were captured. The title, authors, publication year, techniques used, learning method, tasks performed, dataset used, imaging technology used, data pre-processing techniques used, validation technique used, outstanding gaps, and links to codes or external data provided by papers.

RQ1: Types of data and Imaging Modalities used

The first research question aimed at analyzing the type of data used and medical imaging modalities used in training existing MTL models. The study discovered that at least each of the studies considered in the review used some dataset for training. Five (5) studies (9.62%) used primary dataset only, 39 studies (75%) used secondary dataset only, and 8 studies (15.39%) used a combination of primary and secondary datasets. This shows that secondary datasets have the highest level of usage among researchers in this field. This can be attributed to the fact that collecting primary data in the medical imaging domain is both costly and timeconsuming. For instance, Schacky et al., (2021) record that the secondary dataset used consisted of 934 images collected between 2000 and 2020.

In terms of the number of datasets considered in the reviewed articles, we noted that 24 studies (46.15%) used only 1 dataset, 14 studies (26.92%) used 2 datasets, 8 studies (15.38%) used 3 datasets, 4 studies used 4 datasets (7.69%), and only two studies (Butoi et al., 2023), and (Mormont et al., 2021) used above 5 datasets (3.85%). Considering that over 50% of the studies have used more than one dataset, it is evident that for MTL models to be generalizable they need to be exposed to more than one set. However, there is still huge a gap in exploring the generalization abilities of MTL models across multiple datasets since only two studies have gone past five datasets.

The studies reviewed used either a single imaging modality or a combination of several imaging modalities. Figure 2 shows the distribution of medical imaging modalities found in the study. The findings show that MRI is the most commonly used medical imaging modality, followed by CT scan, and then X-ray imaging. The findings also reveal that 23.08% of the study used at least more than one imaging modality. Two studies (Butoi et al., 2023) and (Mormont et al., 2021) have been classified as "Not-specified (Multi-modal) since they have used a combination of datasets that have multiple imaging modalities and they have not categorically stated the imaging modalities used.

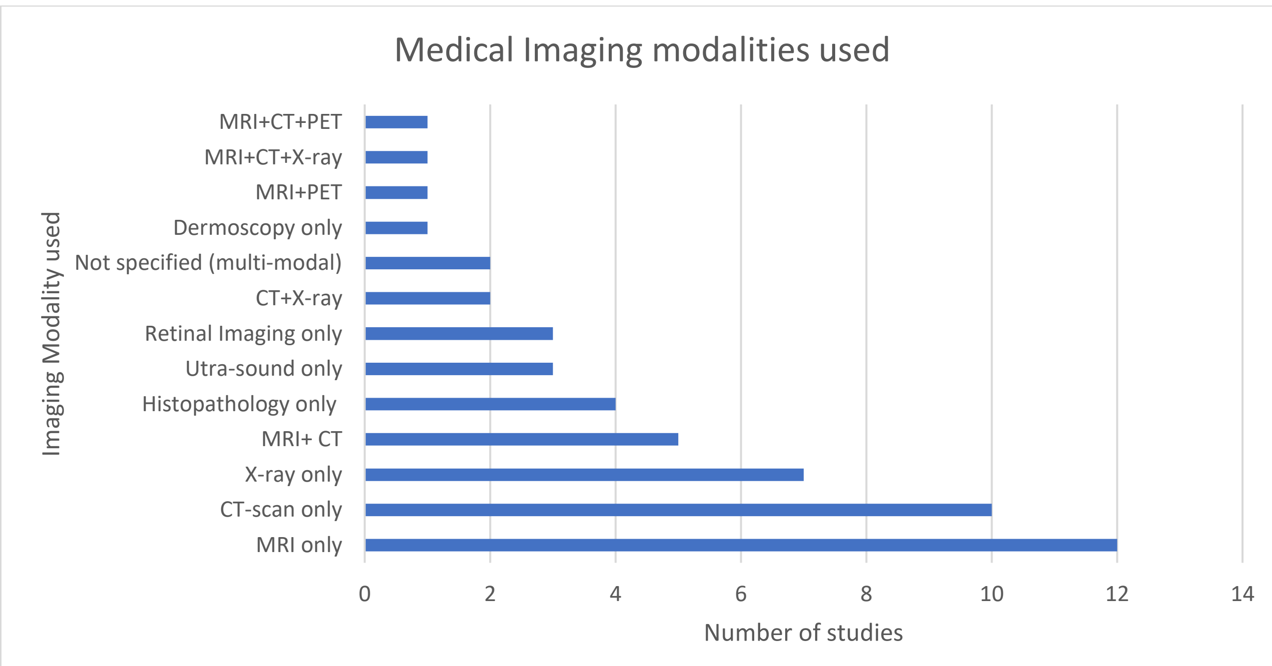

Figure 2. Imaging Modalities Used

RQ2 Multi-Task Learning Techniques used in medical image processing

The second research question aimed at exploring the various techniques that have been used in MTL for medical image processing. This section provides a thorough analysis of the findings.

Learning Methods in MTL

There are four main learning methods for MTL namely hard parameter sharing, soft parameter sharing, multi-task attention network, and cross-task attention network. Hard parameter sharing in MTL involves sharing common layers among the tasks, especially the feature extraction layers, and then the model branches into task-specific layers (Ruder, 2017) as shown in Figure 3. Hard parameter sharing enables a model to find a more generalized representation of all the tasks thus minimizing the risk of overfitting. However, there is a great need to consider the cooperativeness versus the competitiveness of the tasks involved while designing the hard parameter-sharing MTL. This is because hard-parameter sharing is a feature-learning approach that focuses on learning common features among tasks (Zhang & Yang, 2022). Therefore, competing tasks would result in poor performance since the model will struggle to get a global minimum of the competing tasks. Forty-six (46) of the fifty-two articles considered in this review (88.46%) used hard parameter sharing as shown in Table 1.

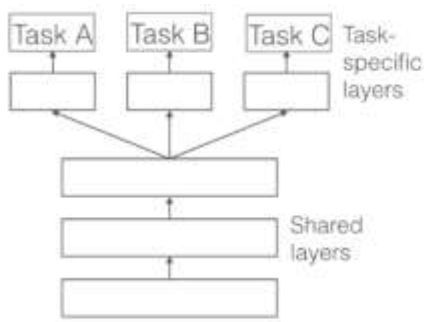

Figure 3.: Hard Parameter Sharing (Ruder, 2017)

Soft parameter sharing involves each task having its model and parameters but knowledge is shared among the tasks through regularization (Ruder, 2017). Common regularization techniques that can be used include L_1 also known as Lasso and L_2 also known as Ridge regularization. L_1 arguments the loss function by adding a penalty term which is equal to the absolute values of the weights of the loss function (Huang et al., 2019). It can be represented mathematically as:

$$Loss\ (L1) = Loss\ (X) + \lambda \sum_i |w_i|. \qquad \textit{Equation 1}$$

L_2 on the other hand argues the loss function by adding a penalty equal to the sum of the squared values of the loss function (Huang et al., 2019). It can be represented mathematically as:

$$Loss\ (L2) = Loss\ (X) + \lambda \sum_i |w_i^2|. \qquad \textit{Equation 2}$$

$Loss\ (X)$ represents the original loss function such as cross-entropy loss or mean squared error, λ represents a hyperparameter that determines the strength of the regularization and the coefficients of the model are represented by w_i. Three articles in this study used soft parameter sharing they represent 5.77% of the total papers considered. Figure 4 shows Soft parameter sharing.

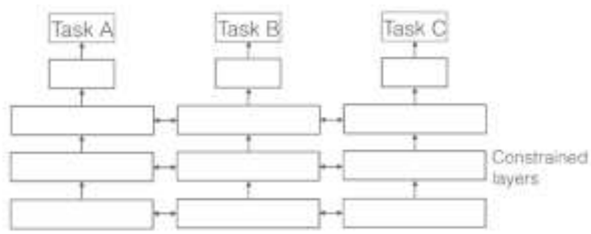

Figure 3. Soft Parameter Sharing(Ruder, 2017)

Multi-task attention network (MTAN) introduced by Liu et al., (2019) consists of shared feature extraction layers but with soft attention mechanisms for each task. The attention mechanism allows for optimizing the learning of task-specific features from the shared feature extractor while still allowing features to be shared across the tasks. From our analysis, no article used MTAN in medical image processing. Kim et al., (2023), argue that although MTAN has gained huge popularity in non-medical image processing due to its ability to analyze pixel-level features, it is challenged in the medical image processing domain since many medical images lack clear pixel-level features. To address this gap Kim et al., (2023) proposed a cross-task attention network (CTAN).

Cross-task Attention Network (CTAN) is a hybrid framework that can process both image-level features and pixel-level features. It consists of a cross-task attention encoder and a cross-task attention bottleneck. The cross-task attention encoder extracts task-specific information in the encoder in a similar way to the attention modules in MTAN. The cross-attention bottleneck captures inter-task interactions across the tasks. Only one article (Kim et al., 2023) used CTAN representing 1.92% of the papers considered in the study. Figure 5 shows CTAN. Two articles among the analyzed articles used hybrids of learning methods. Liu et al., (2017) used a hybrid of soft parameter sharing and hard parameter sharing, while Yang et al., (2021) used a hybrid of soft parameter sharing and CTAN as shown in Table 1.

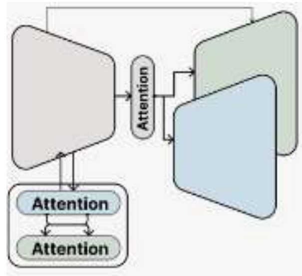

Figure 4. Cross-Task Attention Network (CTAN)(Kim et al., 2023a)

Table 1: MTL learning methods and the studies/articles that have used them.

MTL Learning Method	Studies that use the Method
Hard-Parameter sharing	(Rajpurkar et al., 2017; Gao et al., 2020; Amyar et al., 2020; Chen et al., 2019; Moeskops et al., 2016; Tran et al., 2021; Ngo et al., 2020; Liu et al., 2021; Park et al., 2020; Haque et al., 2021; Song et al., 2020; Zhang et al., 2021; Vesal et al., 2021; Marinov et al., 2023; Pascal et al., 2022; Peng et al., 2019; Haque et al., 2021; Yu et al., 2021; Tan et al., 2018; Cheng et al., 2022; Murugesan et al., 2019; von Schacky et al., 2021; Khosravan & Bagci, 2018; Zhai et al., 2020; Mormont et al., 2021; Eslami et al., 2020; Chen et al., 2019; Gong et al., 2021; Zhou et al., 2020; Hong et al., 2020; Wang et al., 2020; Chamanzar & Nie, 2020; Zhang et al., 2021; Namburete et al., 2018; Kordnoori et al., 2023; Liu et al., 2018; Feng et al., 2018; Vuong et al., 2020; Li et al., 2021; Butoi et al., 2023; K. Zhou et al., 2018; Duan et al., 2019; Sukegawa et al., 2021; Dvornek et al., 2019; He et al., 2021; Yang et al., 2021).
Soft-parameter Sharing	(Alom et al., 2020; Thung et al., 2017; H. Huang et al., 2022)
Cross-task attention Network (CTAN)	(Kim et al., 2023)
Hybrid (Hard-parameter & Soft parameter sharing	(Liu et al., 2017)
Hybrid (Soft parameter sharing & CTAN)	(Yang et al., 2021)

RQ 3 Task Combinations

The third question aimed at analyzing the task combinations that constitute MTL. Different task combinations can constitute the MTL model. Standley et al., (2020) argue that in an MTL setup, it is likely that all tasks can improve in performance if they are cooperating tasks. It is also possible for some tasks to aid others in improving performance while they do not improve in performance. Such tasks are known as loss-guiding tasks or residual tasks. In the field of medical image processing, numerous single tasks can be put in a multi-task setup. They include; segmentation, classification, regression, object detection, and reconstruction, among others.

When all the tasks in the MTL setup fall under a single type of learning such as supervised learning it is called homogenous MTL. On the other hand, if the tasks have combinations of several types of learning such as supervised and semi-supervised learning then it is known as heterogenous MTL(Zhang & Yang, 2022). The scope of a single task varies from one study formulation to another. For instance, in (Butoi et al., 2023) a single task is the segmentation of images from a specific imaging modality like MRI, so their definition of MTL is the ability to jointly segment images from multiple imaging modalities. Others such as (Xue et al., 2021) consider voxel-wise segmentation of pancreas and skeleton extraction from the same CT images dataset as their MTL combination. Mormont et al., (2021) had an MTL setup that combined digital pathology images into a pool with 22 classification tasks and 81 classes.

Therefore, it is evident that the framing of tasks to be included in an MTL model should be informed by whether the included tasks can be split into individual single tasks. Thus, an MTL can be considered to be true if each task can be associated with a loss function. A major notable theme in this study is that for joint learning to happen, the majority of the studies reviewed have formulated a combined loss function that considers the task-specific loss functions of the individual tasks. i.e given three tasks $T_a, T_b, and\ T_c$, and given their loss functions as $L_a, L_{b,} and\ L_c$ then the combined loss can be formulated as:

$$L_{combined} = \lambda_1 L_a + \lambda_2 L_b + \lambda_3 L_c \qquad \textit{Equation 3}$$

Where $\lambda_1, \lambda_2, and\ \lambda_3$ are weights for individual loss function. For instance, take a scenario of an MTL model performing classification and segmentation of medical images from the same dataset. The objective of classification is to assign an input X to a class K. The input space can be defined as $X = (x_1, x_2, x_3, \ldots .x_d) \in \mathbb{R}^d$ for a single input datapoint with d features. This input has labels $y\, y \in \{1,2, \ldots K\}$ where there are K possible classes i.e. a multi-class classification. For binary classification then this would be $y \in \{0, 1\}$. A model $f(X;\theta)$ is formulated as the learning function where θ represents the parameters of the model learned during training. A classification task goes through a probability estimation $P(y = k|X;\theta)$ that an input X belongs to a class given the model parameters θ. Thus, a SoftMax function of this task is represented as:

$$P(y = k|X;\theta) = \frac{\exp\,(f_k(X;\theta))}{\sum_{j=1}^{K} \exp\,(f_j(X;\theta))} \qquad \textit{Equation 4}$$

Where $f_k(X;\theta)$ is the logit (output score) of class k. A loss function such as the cross-entropy loss is used to evaluate how well a model has learned by measuring the difference between predicted labels and actual labels. An average cross-entropy loss for a dataset with N examples can be represented as;

$$l_{cs}(\theta) = -\frac{1}{N} \sum_{i=1}^{N} log\ P(y_i|X_i;\theta) \qquad \textit{Equation 5}$$

given that (xi, yi) is the i[th] training sample. When the $l_{cs}(\theta)$ is too high researchers perform optimization, which is the process of minimizing the loss by adjusting parameters θ. This is mainly done using gradient learning optimizers such as Stochastic gradient descent (SGD), Adam, Adamax, RMSprop, Adagrad, and Adadelta, among others(Dogo et al., 2018).

The objective of segmentation is to assign a label to each pixel in an image. The input space is defined as $X \in \mathbb{R}^{H\times W\times C}$ where X is the input image with features Height (H), Width (W), and RGB Channels (C). The labels for the input image are $Y \in \{1,2, \dots . K\}^{H\times W}$ where there are K possible classes, and each element Y_{ij} represents the class labels of pixels at position (i, j). The model is thus formulated as $F(X; \theta)$ where F is a function that maps the input image X to a matrix of class labels: $F\colon\ \mathbb{R}^{H\times W\times C} \to \{1,2, \dots K\}^{H\times W}$. The probability that a pixel in position (i, j) of image X belongs to k can be computed as $P(Y_{ij} = k|X; \theta)$ Thus a SoftMax function of this task is represented as:

$$P\big(Y_{ij} = k\big|X; \theta\big) = \frac{\exp\,(F_{ij}^{k}(X;\theta))}{\sum_{m=1}^{k} \exp\,(F_{ij}^{m}(X;\theta))} \qquad \textit{Equation 6}$$

Where $F_{ij}^{k}(X; \theta)$ is the logit for class k at pixel (i, j). The average cross entropy loss of a dataset with N images can be represented as:

$$l_{seg}(\theta) = -\frac{1}{N}\sum_{i=1}^{N}\times\sum_{m=1}^{H}\times\sum_{.=1}^{W} log\ P((Y_i)_{mn}|X_i; \theta).$$

Equation 7

Optimization is then applied to reduce the error rate. The combined loss function for the two tasks will be: $L = \lambda_1 l_{cls}(\theta) + \lambda_2 l_{seg}(\theta)$. As the number of tasks continues to increase then the combined loss is adjusted accordingly. For instance in Gao et al.,(Gao et al., 2020) each of the three tasks in the MTL model has a task-specific loss function i.e Tumor detection (bounding box regression loss), segmentation (average cross-entropy loss), and classification (log loss). The three loss functions are then combined to have.

"$L_{uni} = \lambda_1 L_{cls} + \lambda_2 L_{box} + \lambda_3 L_{mask}$" (Gao et al., 2020) *Equation 8*

Network Architectures

Network architecture refers to the designing of a deep neural network used in a particular model. Several architectures have emerged in this review. The choice of architecture was mainly influenced by the tasks being performed by the MTL model. MTL models that had segmentation as part of their task collection used an encoder-decoder architecture. The most prominent encoder-decoder architecture used by majority of the studies that did segmentation is the U-net and its variants (Chen et al., 2019; Ngo et al., 2020; Park et al., 2020; Haque et al., 2021; Vesal et al., 2021; Marinov et al., 2023; Pascal et al., 2022; Murugesan et al., 2019; Alom et al., 2020, Chen et al., 2019; Chamanzar & Nie, 2020; Kordnoori et al., 2023; Butoi et al., 2023; Xue et al., 2021; He et al., 2021).

A U-net is a network architecture that was created primarily for image segmentation. The base U-net has two parts which are the encoder also known as the contracting part and the decoder which is also known as the expanding part. The encoder has a series of blocks with each block having 3*3 convolution followed by RELU(Siddique et al., 2021). The decoder part up samples, concatenates and crops the features. Cropping eliminates edge features since they have the least contextual information(Siddique et al., 2021). The network's energy function can be given as:

$$“E = \sum w(x) \log (Pk_{\ x}(x)) \qquad \textit{Equation 9}$$

Where Pk is the pixel-wise softmax activation function applied over the final feature map” (Siddique et al., 2021).

The other studies that did segmentation used pre-trained CNN architectures (Gao et al., 2020; Kim et al., 2023; Song et al., 2020; Hong et al., 2020; Feng et al., 2018). Some used unique networks such as Cheng et al., (2022) who used a hybrid of U-net and residual network. Huang et al., (2021) who used DeepLab v2. Liu et al., (2021) used a V-net. MTL models that were classified as one of their tasks also used pre-trained CNNs such as (Peng et al., 2019; Vuong et al., 2020). The other studies used custom convolutional neural networks and other types of neural networks to perform their respective tasks. For instance, Dvornek et al., (2019) used recurrent neural networks and long short-term memory. The findings reveal that the choice and design of the network architecture are influenced highly by the task combinations. The network design must be able to meet the objectives of the two tasks, therefore, it should optimize the loss functions of the two tasks and develop a global minimum based on the combined objective functions. This means that moving into the future newer architectures are likely to emerge as more tasks continue to be combined for joint learning.

RQ 5 Open Challenges and Future Research Issues

The last question in this review aimed at evaluating the future of research in Multi-task learning for medical image processing. This section outlines the open challenges in the domain and future research directions. The first open challenge is an expansion of the multi-task scope. This could take different formats. Mormont et al., (2021) have attempted to increase the scope of classification tasks by creating a MTL model with 22 classification tasks. Butoi et al., (2023) have also done some good work in increasing the scope of MTL by having a universal model that can segment medical images from various imaging modalities. Others such as (Rajpurkar et al., 2017), have expanded the scope by creating models that can diagnose multiple 14 disorders though from the same ChestX-ray 14 dataset. However, despite these developments, there is still a need to expand the scope of MTL models to capture; more task combinations, more generalization across different imaging modalities, and even the inclusion of non-imaging tasks such as clinical history as residual tasks to aid MTL image processing. There is also a need to design more optimal MTL models that can achieve high performance at a lesser cost compared to the combined cost of their respective single-task counterparts(Standley et al., 2020).

The issue of task combination is a major challenge that needs to be addressed. Standley et al., (2020) have demonstrated that not all tasks can improve each other when combined. Therefore, researchers need to define a theoretical framework that can define task relatedness. This can be approached either from feature-relatedness or parameter-relatedness (Zhang & Yang, 2018). Soft parameter sharing has been used in most cases when dealing with competing tasks(Ruder, 2017). However, the regularization techniques used still fall short in terms of efficient sharing of parameters among the tasks, thus, resulting in high negative transfer rather than positive transfer(Standley et al., 2020). This is because the ridge and Lasso regularization techniques focus on the use of a penalty to control overfitting. Therefore, there is a need for an advanced regularization technique that can evaluate the quality and impact of shared parameters to minimize negative transfer.

The other open challenge is the development of longitudinal MTL models that integrate seamlessly with clinical workflows. This means that there is a need for MTL models that can monitor patient medical image data over some time to assess the progress of the medical condition under consideration. The next challenge is the need for more accurate generative MTL models. There are numerous medical conditions such as cancers that can mutate, thus a generative MTL model can play a critical role in generating future mutation images of such medical conditions. There are also cases where organs could be severely damaged. A generative model can help in regenerating images of the original organ for purposes of damage assessment.

CONCLUSION

Deep learning models are increasingly improving in their capabilities with Multi-task learning attracting the attention of many researchers. Multi-task learning has demonstrated its capability to provide more generalizable and cost-effective models. This is especially critical in fields such as medical imaging where single-task models are not only inefficient but also costly. This inefficiency and high cost of single-task models are mainly caused by the lack of large-scale medical datasets such as ImageNet. Also, the distinct differences between natural images contained in ImageNet and medical images make transfer learning not an optimal solution. This, therefore, places multi-task learning at the best place in processing medical images. Through this study, we have analyzed 52 primary papers and presented our findings to answer the research questions. In the first RQ question, we present a detailed analysis of imaging modalities and types of data used in medical image processing. In the second question, we have analyzed the key MTL techniques that can be used in both computer vision and non-computer vision tasks. In our third research question, we have critically analyzed how task combinations can be formulated in medical image processing. Finally, we provide an analysis of outstanding gaps that future researchers can work on addressing. Addressing the identified gaps would be a great path toward artificial general intelligence in medical image processing.

IMPLICATIONS

The review evaluates the current state of research in medical image processing using multi-task learning (MTL), offering readers a comprehensive understanding of both the theoretical and practical aspects of the field. Theoretically, it provides an in-depth analysis of cutting-edge techniques in task combinations, learning methods, formulation of joint learning functions, and architectural design of MTL networks. Practically, it critically examines the training of MTL models with various datasets and medical imaging modalities. Zhang and Yang, (2018) argue that a well-designed MTL setup can theoretically and practically improve the overall performance of the cooperating task. This has been validated by (Amyar et al., 2020), who did joint segmentation of CT images and detection of COVID-19 from the images and achieved a dice coefficient of 0.88 specificity of 0.97, sensitivity of 0.90 for the segmentation task and area under the curve (AUC) of 97% for classification.

Butoi et al., (2023) have also demonstrated that multi-modal fusion in medical image segmentation can improve the generalizability of models. Additionally, the review identifies open gaps in the field, guiding researchers and practitioners toward future research directions to advance the discipline. The study highlights the high cost of collecting primary datasets, which hinders the development of MTL models, particularly for emerging medical disorders in many parts of the world. This finding calls for policymakers in healthcare to create policies that encourage the ethical collection and sharing of medical imaging datasets, facilitating the customization of MTL solutions for specific disorders. Federated learning is one approach that can guarantee ethical sharing of data between medical facilities and researchers especially when researchers need data from multiple facilities. Also, there is a need for researchers to enhance the explainability of their models to ensure they are not biased as part of enhancing ethics in this field.

DECLARATION

Conflict of Interest

The author declared that there is no conflict of interest.

Inform Consent

This may not be applicable because this is a review article, and respondents are not involved.

Ethics Approval

It is not applicable because this is a review article, and no respondents are required.

REFERENCES

Alom, M. Z., Aspiras, T., Taha, T. M., Bowen, T., & Asari, V. K. (2020). MitosisNet: End-to-End Mitotic Cell Detection by Multi-Task Learning. *IEEE Access*, *8*, 68695–68710. IEEE Access. https://doi.org/10.1109/ACCESS.2020.2983995

Amyar, A., Modzelewski, R., Li, H., & Ruan, S. (2020). Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation. *Computers in Biology and Medicine*, *126*, 104037. https://doi.org/10.1016/j.compbiomed.2020.104037

Asgari Taghanaki, S., Abhishek, K., Cohen, J. P., Cohen-Adad, J., & Hamarneh, G. (2021). Deep semantic segmentation of natural and medical images: A review. *Artificial Intelligence Review*, *54*(1), 137–178. https://doi.org/10.1007/s10462-020-09854-1

Buettner, R., Bilo, M., Bay, N., & Zubac, T. (2020). A Systematic Literature Review of Medical Image Analysis Using Deep Learning. *2020 IEEE Symposium on Industrial Electronics & Applications (ISIEA)*, 1–4. https://doi.org/10.1109/ISIEA49364.2020.9188131

Butoi, V. I., Ortiz, J. J. G., Ma, T., Sabuncu, M. R., Guttag, J., & Dalca, A. V. (2023). *UniverSeg: Universal Medical Image Segmentation*. 21438–21451. https://openaccess.thecvf.com/content/ICCV2023/html/Butoi_UniverSeg_Universal_Medical_Image_Segmentation_ICCV_2023_paper.html

Çetiner, H., & Metlek, S. (2023). DenseUNet+: A novel hybrid segmentation approach based on multi-modality images for brain tumor segmentation. *Journal of King Saud University - Computer and Information Sciences*, *35*(8), 101663. https://doi.org/10.1016/j.jksuci.2023.101663

Chamanzar, A., & Nie, Y. (2020). Weakly Supervised Multi-Task Learning for Cell Detection and Segmentation. *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 513–516. https://doi.org/10.1109/ISBI45749.2020.9098518

Chen, C., Bai, W., & Rueckert, D. (2019). Multi-task Learning for Left Atrial Segmentation on GE-MRI. In M. Pop, M. Sermesant, J. Zhao, S. Li, K. McLeod, A. Young, K. Rhode, & T. Mansi (Eds.), *Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges* (pp. 292–301). Springer International Publishing. https://doi.org/10.1007/978-3-030-12029-0_32

Chen, S., Bortsova, G., García-Uceda Juárez, A., van Tulder, G., & de Bruijne, M. (2019). Multi-task Attention-Based Semi-supervised Learning for Medical Image Segmentation. In D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, & A. Khan (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019* (pp. 457–465). Springer International Publishing. https://doi.org/10.1007/978-3-030-32248-9_51

Cheng, J., Liu, J., Kuang, H., & Wang, J. (2022a). A Fully Automated Multimodal MRI-Based Multi-Task Learning for Glioma Segmentation and IDH Genotyping. *IEEE Transactions on Medical Imaging*, *41*(6), 1520–1532. IEEE Transactions on Medical Imaging. https://doi.org/10.1109/TMI.2022.3142321

Cheng, J., Liu, J., Kuang, H., & Wang, J. (2022b). A Fully Automated Multimodal MRI-Based Multi-Task Learning for Glioma Segmentation and IDH Genotyping. *IEEE Transactions on Medical Imaging*, *41*(6), 1520–1532. IEEE Transactions on Medical Imaging. https://doi.org/10.1109/TMI.2022.3142321

D, V., & Thenmozhi, R. (2023). A Systematic Literature Review on Deep Learning Based Medical Image Segmentation. *International Journal of Intelligent Systems and Applications in Engineering*, *11*(5s), Article 5s.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

Dogo, E. M., Afolabi, O. J., Nwulu, N. I., Twala, B., & Aigbavboa, C. O. (2018). A Comparative Analysis of Gradient Descent-Based Optimization Algorithms on Convolutional Neural Networks. *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, 92–99. https://doi.org/10.1109/CTEMS.2018.8769211

Duan, J., Bello, G., Schlemper, J., Bai, W., Dawes, T. J. W., Biffi, C., de Marvao, A., Doumoud, G., O'Regan, D. P., & Rueckert, D. (2019). Automatic 3D Bi-Ventricular Segmentation of Cardiac Images by a Shape-Refined Multi-Task Deep Learning Approach. *IEEE Transactions on Medical Imaging*, *38*(9), 2151–2164. IEEE Transactions on Medical Imaging. https://doi.org/10.1109/TMI.2019.2894322

Dvornek, N. C., Li, X., Zhuang, J., & Duncan, J. S. (2019). Jointly Discriminative and Generative Recurrent Neural Networks for Learning from fMRI. *Machine Learning in Medical Imaging. MLMI (Workshop)*, *11861*, 382–390. https://doi.org/10.1007/978-3-030-32692-0_44

Eslami, M., Tabarestani, S., Albarqouni, S., Adeli, E., Navab, N., & Adjouadi, M. (2020). Image-to-images translation for Multi-Task Organ Segmentation and Bone Suppression in Chest X-Ray Radiography. *IEEE Transactions on Medical Imaging*, *39*(7), 2553–2565. IEEE Transactions on Medical Imaging. https://doi.org/10.1109/TMI.2020.2974159

Feng, Z., Nie, D., Wang, L., & Shen, D. (2018). Semi-supervised learning for pelvic MR image segmentation based on multi-task residual fully convolutional networks. *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 885–888. https://doi.org/10.1109/ISBI.2018.8363713

Gao, F., Yoon, H., Wu, T., & Chu, X. (2020). A feature transfer enabled multi-task deep learning model on medical imaging. *Expert Systems with Applications*, *143*, 112957. https://doi.org/10.1016/j.eswa.2019.112957

Gong, H., Chen, G., Wang, R., Xie, X., Mao, M., Yu, Y., Chen, F., & Li, G. (2021). Multi-Task Learning For Thyroid Nodule Segmentation With Thyroid Region Prior. *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 257–261. https://doi.org/10.1109/ISBI48211.2021.9434087

Haque, A., Imran, A.-A.-Z., Wang, A., & Terzopoulos, D. (2021a). Generalized Multi-Task Learning from Substantially Unlabeled Multi-Source Medical Image Data. *Machine Learning for Biomedical Imaging*, *1*(October 2021), 1–25. https://doi.org/10.59275/j.melba.2021-d8a3

He, K., Lian, C., Adeli, E., Huo, J., Gao, Y., Zhang, B., Zhang, J., & Shen, D. (2021). MetricUNet: Synergistic image- and voxel-level learning for precise prostate segmentation via online sampling. *Medical Image Analysis*, *71*, 102039. https://doi.org/10.1016/j.media.2021.102039

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Helland, R. H., Ferles, A., Pedersen, A., Kommers, I., Ardon, H., Barkhof, F., Bello, L., Berger, M. S., Dunås, T., Nibali, M. C., Furtner, J., Hervey-Jumper, S., Idema, A. J. S., Kiesel, B., Tewari, R. N., Mandonnet, E., Müller, D. M. J., Robe, P. A., Rossi, M., … Bouget, D. (2023). Segmentation of glioblastomas in early post-operative multi-modal MRI with deep neural networks. *Scientific Reports*, *13*(1), 18897. https://doi.org/10.1038/s41598-023-45456-x

Hong, Y., Wei, B., Han, Z., Li, X., Zheng, Y., & Li, S. (2020). MMCL-Net: Spinal disease diagnosis in global mode using progressive multi-task joint learning. *Neurocomputing*, *399*, 307–316. https://doi.org/10.1016/j.neucom.2020.01.112

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications* (No. arXiv:1704.04861). arXiv. http://arxiv.org/abs/1704.04861

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2261–2269. https://doi.org/10.1109/CVPR.2017.243

Huang, H., Chen, Q., Lin, L., Cai, M., Zhang, Q., Iwamoto, Y., Han, X., Furukawa, A., Kanasaki, S., Chen, Y.-W., Tong, R., & Hu, H. (2022). MTL-ABS3Net: Atlas-Based Semi-Supervised Organ Segmentation Network With Multi-Task Learning for Medical Images. *IEEE Journal of Biomedical and Health Informatics*, *26*(8), 3988–3998. IEEE Journal of Biomedical and Health Informatics. https://doi.org/10.1109/JBHI.2022.3153406

Huang, H., Yang, G., Zhang, W., Xu, X., Yang, W., Jiang, W., & Lai, X. (2021). A Deep Multi-Task Learning Framework for Brain Tumor Segmentation. *Frontiers in Oncology*, *11*. https://doi.org/10.3389/fonc.2021.690244

Huang, Y., Schell, C., Huber, T. B., Şimşek, A. N., Hersch, N., Merkel, R., Gompper, G., & Sabass, B. (2019). Traction force microscopy with optimized regularization and automated Bayesian parameter selection for comparing cells. *Scientific Reports*, *9*(1), 539. https://doi.org/10.1038/s41598-018-36896-x

Khosravan, N., & Bagci, U. (2018). Semi-Supervised Multi-Task Learning for Lung Cancer Diagnosis. *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 710–713. https://doi.org/10.1109/EMBC.2018.8512294

Kim, S., Purdie, T. G., & McIntosh, C. (2023a). Cross-Task Attention Network: Improving Multi-task Learning for Medical Imaging Applications. In M. E. Celebi, M. S. Salekin, H. Kim, S. Albarqouni, C. Barata, A. Halpern, P. Tschandl, M. Combalia, Y. Liu, G. Zamzmi, J. Levy, H. Rangwala, A. Reinke, D. Wynn, B. Landman, W.-K. Jeong, Y. Shen, Z. Deng, S. Bakas, … D. Xu (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023 Workshops* (pp. 119–128). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-47401-9_12

Kim, S., Purdie, T. G., & McIntosh, C. (2023b). Cross-Task Attention Network: Improving Multi-task Learning for Medical Imaging Applications. In M. E. Celebi, M. S. Salekin, H. Kim, S. Albarqouni, C. Barata, A. Halpern, P. Tschandl, M. Combalia, Y. Liu, G. Zamzmi, J. Levy, H. Rangwala, A. Reinke, D. Wynn, B. Landman, W.-K. Jeong, Y. Shen, Z. Deng, S.

Bakas, … D. Xu (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023 Workshops* (pp. 119–128). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-47401-9_12

Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, *51*(1), 7–15. https://doi.org/10.1016/j.infsof.2008.09.009

Kordnoori, S., Sabeti, M., Mostafaei, H., & Banihashemi, S. S. A. (2023). Analysis of lung scan imaging using deep multi-task learning structure for COVID-19 disease. *IET Image Processing*, *17*(5), 1534–1545. https://doi.org/10.1049/ipr2.12736

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, *25*. https://proceedings.neurips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c843 6e924a68c45b-Abstract.html

Li, J., Zhao, G., Tao, Y., Zhai, P., Chen, H., He, H., & Cai, T. (2021). Multi-task contrastive learning for automatic CT and X-ray diagnosis of COVID-19. *Pattern Recognition*, *114*, 107848. https://doi.org/10.1016/j.patcog.2021.107848

Liu, L., Dou, Q., Chen, H., Olatunji, I. E., Qin, J., & Heng, P.-A. (2018). MTMR-Net: Multi-task Deep Learning with Margin Ranking Loss for Lung Nodule Analysis. In D. Stoyanov, Z. Taylor, G. Carneiro, T. Syeda-Mahmood, A. Martel, L. Maier-Hein, J. M. R. S. Tavares, A. Bradley, J. P. Papa, V. Belagiannis, J. C. Nascimento, Z. Lu, S. Conjeti, M. Moradi, H. Greenspan, & A. Madabhushi (Eds.), *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (pp. 74–82). Springer International Publishing. https://doi.org/10.1007/978-3-030-00889-5_9

Liu, M., Zhang, J., Adeli, E., & Shen, D. (2017). Deep Multi-task Multi-channel Learning for Joint Classification and Regression of Brain Status. In M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. L. Collins, & S. Duchesne (Eds.), *Medical Image Computing and Computer Assisted Intervention − MICCAI 2017* (pp. 3–11). Springer International Publishing. https://doi.org/10.1007/978-3-319-66179-7_1

Liu, S., Johns, E., & Davison, A. J. (2019). *End-To-End Multi-Task Learning With Attention*. 1871–1880. https://openaccess.thecvf.com/content_CVPR_2019/html/Liu_End-To-End_Multi-Task_Learning_With_Attention_CVPR_2019_paper.html

Liu, S., Li, Y., Li, X., & Cao, G. (2021). Shape-aware Multi-task Learning for Semi-supervised 3D Medical Image Segmentation. *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 1418–1423. https://doi.org/10.1109/BIBM52615.2021.9669523

Marinov, Z., Reiß, S., Kersting, D., Kleesiek, J., & Stiefelhagen, R. (2023). *Mirror U-Net: Marrying Multimodal Fission with Multi-Task Learning for Semantic Segmentation in Medical Imaging*. 2283–2293. https://openaccess.thecvf.com/content/ICCV2023W/CVAMD/html/Marinov_Mirror_U-Net_Marrying_Multimodal_Fission_with_Multi-Task_Learning_for_Semantic_ICCVW_2023_paper.html

Moeskops, P., Wolterink, J. M., van der Velden, B. H. M., Gilhuijs, K. G. A., Leiner, T., Viergever, M. A., & Išgum, I. (2016). Deep Learning for Multi-task Medical Image Segmentation in Multiple Modalities. In S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, & W. Wells

(Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016* (pp. 478–486). Springer International Publishing. https://doi.org/10.1007/978-3-319-46723-8_55

Mormont, R., Geurts, P., & Marée, R. (2021). Multi-Task Pre-Training of Deep Neural Networks for Digital Pathology. *IEEE Journal of Biomedical and Health Informatics*, *25*(2), 412–421. IEEE Journal of Biomedical and Health Informatics. https://doi.org/10.1109/JBHI.2020.2992878

Murugesan, B., Sarveswaran, K., Shankaranarayana, S. M., Ram, K., Joseph, J., & Sivaprakasam, M. (2019). Psi-Net: Shape and boundary-aware joint multi-task deep network for medical image segmentation. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 7223–7226. https://doi.org/10.1109/EMBC.2019.8857339

Namburete, A. I. L., Xie, W., Yaqub, M., Zisserman, A., & Noble, J. A. (2018). Fully automated alignment of 3D fetal brain ultrasound to a canonical reference space using multi-task learning. *Medical Image Analysis*, *46*, 1–14. https://doi.org/10.1016/j.media.2018.02.006

Ngo, D.-K., Tran, M.-T., Kim, S.-H., Yang, H.-J., & Lee, G.-S. (2020). Multi-Task Learning for Small Brain Tumor Segmentation from MRI. *Applied Sciences*, *10*(21), Article 21. https://doi.org/10.3390/app10217790

Park, S., Jeong, W., & Moon, Y. S. (2020). X-ray Image Segmentation using Multi-task Learning. *KSII Transactions on Internet and Information Systems (TIIS)*, *14*(3), 1104–1120. https://doi.org/10.3837/tiis.2020.03.011

Pascal, L., Perdomo, O. J., Bost, X., Huet, B., Otálora, S., & Zuluaga, M. A. (2022). Multi-task deep learning for glaucoma detection from color fundus images. *Scientific Reports*, *12*(1), 12361. https://doi.org/10.1038/s41598-022-16262-8

Peng, T., Boxberg, M., Weichert, W., Navab, N., & Marr, C. (2019). Multi-task Learning of a Deep K-Nearest Neighbour Network for Histopathological Image Classification and Retrieval. In D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, & A. Khan (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019* (pp. 676–684). Springer International Publishing. https://doi.org/10.1007/978-3-030-32239-7_75

Qummar, S., Khan, F. G., Shah, S., Khan, A., Shamshirband, S., Rehman, Z. U., Khan, I. A., & Jadoon, W. (2019). A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection. *IEEE Access*, *7*, 150530–150539. https://doi.org/10.1109/ACCESS.2019.2947484

Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M. P., & Ng, A. Y. (2017). *CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning* (No. arXiv:1711.05225). arXiv. http://arxiv.org/abs/1711.05225

Ruder, S. (2017). *An Overview of Multi-Task Learning in Deep Neural Networks* (No. arXiv:1706.05098). arXiv. http://arxiv.org/abs/1706.05098

Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V. (2021). U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. *IEEE Access*, *9*, 82031–82057. IEEE Access. https://doi.org/10.1109/ACCESS.2021.3086020

Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition* (No. arXiv:1409.1556). arXiv. http://arxiv.org/abs/1409.1556

Sohail, A., Nawaz, N. A., Shah, A. A., Rasheed, S., Ilyas, S., & Ehsan, M. K. (2022). A Systematic Literature Review on Machine Learning and Deep Learning Methods for Semantic Segmentation. *IEEE Access*, *10*, 134557–134570. IEEE Access. https://doi.org/10.1109/ACCESS.2022.3230983

Song, L., Lin, J., Wang, Z. J., & Wang, H. (2020a). An End-to-End Multi-Task Deep Learning Framework for Skin Lesion Analysis. *IEEE Journal of Biomedical and Health Informatics*, *24*(10), 2912–2921. https://doi.org/10.1109/JBHI.2020.2973614

Song, L., Lin, J., Wang, Z. J., & Wang, H. (2020b). An End-to-End Multi-Task Deep Learning Framework for Skin Lesion Analysis. *IEEE Journal of Biomedical and Health Informatics*, *24*(10), 2912–2921. IEEE Journal of Biomedical and Health Informatics. https://doi.org/10.1109/JBHI.2020.2973614

Standley, T., Zamir, A., Chen, D., Guibas, L., Malik, J., & Savarese, S. (2020). Which Tasks Should Be Learned Together in Multi-task Learning? *Proceedings of the 37th International Conference on Machine Learning*, 9120–9132. https://proceedings.mlr.press/v119/standley20a.html

Sukegawa, S., Yoshii, K., Hara, T., Matsuyama, T., Yamashita, K., Nakano, K., Takabatake, K., Kawai, H., Nagatsuka, H., & Furuki, Y. (2021). Multi-Task Deep Learning Model for Classification of Dental Implant Brand and Treatment Stage Using Dental Panoramic Radiograph Images. *Biomolecules*, *11*(6), Article 6. https://doi.org/10.3390/biom11060815

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1–9. https://doi.org/10.1109/CVPR.2015.7298594

Tan, C., Zhao, L., Yan, Z., Li, K., Metaxas, D., & Zhan, Y. (2018). Deep multi-task and task-specific feature learning network for robust shape-preserved organ segmentation. *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 1221–1224. https://doi.org/10.1109/ISBI.2018.8363791

Tan, M., & Le, Q. V. (2020). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks* (No. arXiv:1905.11946). arXiv. http://arxiv.org/abs/1905.11946

Thung, K.-H., Yap, P.-T., & Shen, D. (2017). Multi-stage Diagnosis of Alzheimer's Disease with Incomplete Multimodal Data via Multi-task Deep Learning. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support : Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC.*, *10553*, 160–168. https://doi.org/10.1007/978-3-319-67558-9_19

Tran, M.-T., Kim, S.-H., Yang, H.-J., & Lee, G.-S. (2021). Multi-Task Learning for Medical Image Inpainting Based on Organ Boundary Awareness. *Applied Sciences*, *11*(9), Article 9. https://doi.org/10.3390/app11094247

Vesal, S., Gu, M., Maier, A., & Ravikumar, N. (2021). Spatio-Temporal Multi-Task Learning for Cardiac MRI Left Ventricle Quantification. *IEEE Journal of Biomedical and Health Informatics*, *25*(7), 2698–2709. IEEE Journal of Biomedical and Health Informatics. https://doi.org/10.1109/JBHI.2020.3046449 von Schacky, C. E., Wilhelm, N. J., Schäfer, V. S., Leonhardt, Y., Gassert, F. G., Foreman, S. C., Gassert, F. T., Jung, M., Jungmann, P. M., Russe, M. F., Mogler, C., Knebel, C., von Eisenhart-Rothe, R., Makowski, M. R., Woertler, K., Burgkart, R., & Gersing, A. S. (2021). Multitask Deep Learning for Segmentation and Classification of Primary Bone Tumors on Radiographs. *Radiology*, *301*(2), 398–406. https://doi.org/10.1148/radiol.2021204531

Vuong, T. L. T., Lee, D., Kwak, J. T., & Kim, K. (2020). Multi-task Deep Learning for Colon Cancer Grading. *2020 International Conference on Electronics, Information, and Communication (ICEIC)*, 1–2. https://doi.org/10.1109/ICEIC49074.2020.9051305

Wang, G., Gong, E., Banerjee, S., Martin, D., Tong, E., Choi, J., Chen, H., Wintermark, M., Pauly, J. M., & Zaharchuk, G. (2020). Synthesize High-Quality Multi-Contrast Magnetic Resonance Imaging From Multi-Echo Acquisition Using Multi-Task Deep Generative Model. *IEEE Transactions on Medical Imaging*, *39*(10), 3089–3099. IEEE Transactions on Medical Imaging. https://doi.org/10.1109/TMI.2020.2987026

Xue, J., He, K., Nie, D., Adeli, E., Shi, Z., Lee, S.-W., Zheng, Y., Liu, X., Li, D., & Shen, D. (2021). Cascaded MultiTask 3-D Fully Convolutional Networks for Pancreas Segmentation. *IEEE Transactions on Cybernetics*, *51*(4), 2153–2165. IEEE Transactions on Cybernetics. https://doi.org/10.1109/TCYB.2019.2955178

Yang, Z., Cao, Z., Zhang, Y., Tang, Y., Lin, X., Ouyang, R., Wu, M., Han, M., Xiao, J., Huang, L., Wu, S., Chang, P., & Ma, J. (2021). MommiNet-v2: Mammographic multi-view mass identification networks. *Medical Image Analysis*, *73*, 102204. https://doi.org/10.1016/j.media.2021.102204

Yu, C., Gao, Z., Zhang, W., Yang, G., Zhao, S., Zhang, H., Zhang, Y., & Li, S. (2021). Multitask Learning for Estimating Multitype Cardiac Indices in MRI and CT Based on Adversarial Reverse Mapping. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(2), 493–506. IEEE Transactions on Neural Networks and Learning Systems. https://doi.org/10.1109/TNNLS.2020.2984955

Zhai, P., Tao, Y., Chen, H., Cai, T., & Li, J. (2020). Multi-Task Learning for Lung Nodule Classification on Chest CT. *IEEE Access*, *8*, 180317–180327. IEEE Access. https://doi.org/10.1109/ACCESS.2020.3027812

Zhang, B., Liu, H., Luo, H., & Li, K. (2021). Automatic quality assessment for 2D fetal sonographic standard plane based on multitask learning. *Medicine*, *100*(4), e24427. https://doi.org/10.1097/MD.0000000000024427

Zhang, Y., Li, H., Du, J., Qin, J., Wang, T., Chen, Y., Liu, B., Gao, W., Ma, G., & Lei, B. (2021). 3D Multi-Attention Guided Multi-Task Learning Network for Automatic Gastric Tumor Segmentation and Lymph Node Classification. *IEEE Transactions on Medical Imaging*, *40*(6), 1618–1631. *IEEE Transactions on Medical Imaging*, *40*(6), 1618-1631. https://doi.org/10.1109/TMI.2021.3062902

Zhang, Y., & Yang, Q. (2018). An overview of multi-task learning. *National Science Review*, *5*(1), 30–43. https://doi.org/10.1093/nsr/nwx105

Zhang, Y., & Yang, Q. (2022). A Survey on Multi-Task Learning. *IEEE Transactions on Knowledge and Data Engineering*, *34*(12), 5586–5609. IEEE Transactions on Knowledge and Data Engineering. https://doi.org/10.1109/TKDE.2021.3070203

Zhou, C., Ding, C., Wang, X., Lu, Z., & Tao, D. (2020). One-Pass Multi-Task Networks With Cross-Task Guided Attention for Brain Tumor Segmentation. *IEEE Transactions on Image Processing*, 29, 4516–4529. IEEE Transactions on Image Processing. https://doi.org/10.1109/TIP.2020.2973510

Zhou, K., Gu, Z., Liu, W., Luo, W., Cheng, J., Gao, S., & Liu, J. (2018). Multi-Cell Multi-Task Convolutional Neural Networks for Diabetic Retinopathy Grading. *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2724–2727. https://doi.org/10.1109/EMBC.2018.8512828

Author's Biography

Jackson Kamiri is a Tutorial Fellow and a Ph.D. Information Technology student in the Department of Information Technology, Murang'a University of Technology. He holds an MSc. in Information Technology from Murang'a University of Technology, Kenya. His research interests include deep learning, Image processing, and Big data analytics.

Geoffrey Mariga Wambugu is a senior lecturer in the Department of Information Technology, at Murang'a University of Technology. He holds a Ph.D. in Information Technology from Jomo Kenyatta University of Agriculture and Technology, Kenya. His research interests include Probabilistic Machine Learning, Natural Language Processing, Data mining, Deep learning, and Big Data Analytics.

Aaron Mogeni Oirere is a lecturer in the Department of Computer Science, at the Murang'a University of Technology. He holds a Ph.D. in Computer Science from Dr. Babasaheb Ambedkar Marathwada University, India. His research interests include Deep Learning, Text Analytics, Speech recognition, Human-computer Interaction, and Image processing.