

# LEVERAGING DEEP LEARNING FOR COMPREHENSIVE CLASSIFICATION OF RENAL DISEASES: A TRANSFER LEARNING APPROACH

**Anonymous authors**

Paper under double-blind review

## ABSTRACT

The nightmare of cancer as a leading cause of premature deaths worldwide is becoming real and turns out to be one of the major problems of humanity nowadays. Cancer diagnostics at the early stage is Critical to cancer recovery and survival. In this context, renal diseases, including kidney cysts, stones, and tumors, pose significant global health challenges, affecting approximately 12% of the population and contributing to chronic kidney disease (CKD). Notably, renal cancer ranks as the tenth most prevalent cancer type, accounting for 2.7% of all cancer cases. This work presents a deep learning (DL) framework utilizing transfer learning (TL) for the early detection of renal diseases and categorizing the conditions into four binary classifications: Cyst\_vs\_Normal, Cyst\_vs\_Stone, Cyst\_vs\_Tumor, and Stone\_vs\_Tumor, allowing for a more specific understanding of each stage. By analyzing CT scans and microscopic histopathology images, the framework employs convolutional neural networks (CNNs) with pre-trained models to facilitate automatic and precise classification of renal conditions. Specifically, two CNN models ResNet-50 and EfficientNetV2 are implemented, providing a comprehensive analysis of each stage of the DL architecture. Comparative evaluations of training outcomes across various datasets revealed that EfficientNetV2 performed marginally better than ResNet-50, achieving an impressive testing accuracy of up to 100% for all cases. These results underscore the effectiveness of the DL-based system and highlight its potential for widespread clinical application in renal disease diagnosis.

Keywords- CNN, kidney, image classification, deep learning, transfer learning

## 1 INTRODUCTION

Renal organs are vital, bean-shaped located below the rib cage that filters blood, removes waste, and balances fluids and electrolytes, with each kidney containing about a million nephrons Raghavendra & Vidya (2013). Additionally, they control blood pressure, stimulate red blood cell production through hormone secretion, and maintain overall homeostasis, making their proper function essential for health. Renal cancer originates in the kidneys when malignant cells form in the tubules, often requiring treatments like immunotherapy Navani & Heng (2023). Early detection and advanced medical interventions can significantly improve outcomes for affected individuals Jacobson (2013). Kidney cancer is a growing public health concern, in the year 2022, it ranked as the 14th most common cancer and the 16th leading cause of cancer-related death globally Can (2022) to the 5th by 2040 Foreman et al. (2018). Kidney diseases commonly consist of cysts, stones, and renal cell carcinoma (RCC) Hsieh et al. (2017), while nephrolithiasis affects approximately 12% of the global population Alelign & Petros (2018). Despite control efforts, the prevalence of these conditions continues to increase, highlighting the need for enhanced medical interventions and public health strategies Hsieh et al. (2017). Computed tomography (CT) scans are particularly effective for kidney examinations, offering three-dimensional, cross-sectional images ideal for identifying abnormalities like cysts, stones, and tumors Sagel et al. (1977).

It is challenging for doctors to clinically determine invasive cancer by identifying the captured images, and it needs to administer safe and expensive treatments for the patients Wang et al. (2020). Therefore, tracking the growth of every stage is essential to develop customized medications based

054 on a patient’s disease profile Hsieh et al. (2017). There is a significant global shortage of nephrol-  
055 ogists and radiologists, particularly in Asia, where there is only about one nephrologist for every  
056 million people, in contrast to Europe, which has approximately 25.3 nephrologists per million Sozio  
057 et al. (2021). Given the widespread impact of kidney diseases and the scarcity of specialists, devel-  
058 oping DL-based models to assist in detecting kidney abnormalities has become essential Bi et al.  
059 (2022). Recent advancements in DL-based models for vision tasks offer promising solutions to  
060 support doctors and alleviate patient suffering.

061 This work is organized as follows: Section 2 discusses recent work that has been done previously.  
062 Section 3 provides an overview of the dataset and details of the model training process, including  
063 the initialization of weights. Additionally, this section offers a comprehensive overview of the DL  
064 architectures employed. Two different CNN models are introduced, evaluating their performance  
065 for detection across the four stages of renal disease. Finally, Section 4 presents concluding remarks  
066 and outlines potential pathways for future works.

## 070 2 RELATED WORKS

073 CNNs are a powerful DL algorithm commonly used for classifying grid-like data, such as images Si-  
074 monyan & Zisserman (2015); Szegedy et al. (2014); He et al. (2015a); Tan & Le (2020). Specifically,  
075 renal US images are enhanced using median and Gaussian filtering techniques and morphological  
076 operations Verma et al. (2017). Relevant features from images are extracted using various unsu-  
077 pervised techniques and classified using supervised algorithms. In Aksakalli et al. (2021), authors  
078 demonstrated a range of traditional supervised machine learning algorithms, including decision tree  
079 (DT), random forests (RF), K-nearest neighbors (KNN), and multilayer perceptron (MLP), as well as  
080 CNN. They achieved the best F1 score of 0.853 with those methods. In Sudharson & Kokil (2020),  
081 they employed pre-trained CNN models like ResNet-101, MobileNet-v2, and ShuffleNet to extract  
082 features from kidney US images, got an accuracy of 95.58% using support vector machine (SVM).  
083 In Fu et al. (2021) residual dual-attention (RDA) module utilized for segmenting kidney cysts from  
084 the CT images. In Zheng et al. (2019), combined features extracted through TL approaches, which  
085 were subsequently utilized to distinguish between affected and non-affected ultrasound images by  
086 SVM classifier. In Parakh et al. (2019), two consecutive CNN models were employed: first CNN  
087 identified the urinary region, and second CNN detected the existence of stones, both achieved an ac-  
088 curacy of 95%. In Yildirim et al. (2021) introduced an automated method for the detection of kidney  
089 stones using coronal CT images and DL techniques, achieving an accuracy of 96.82%. In Blau et al.  
090 (2018), researchers developed a system to detect kidney cysts in the images of abdominal CT scans  
091 by utilizing a fully connected CNN. They reported an 84.3% true-positive rate (TPR) for their ap-  
092 proach. In recent studies on kidney disease detection from CT images, the EANet, ResNet50, and a  
093 customized CNN model achieved accuracies of 83.65%, 87.92%, and 98.66%, respectively Hossain  
094 et al. (2023). In summary, initiatives employing ML and DL approaches to classify kidney-related  
095 radiological findings have demonstrated encouraging outcomes, primarily concentrating on CT and  
096 US images.

096 The rapid advancement of CNNs has led to the development and utilization of various DL architec-  
097 tures Khan et al. (2020), including EfficientNet Tan & Le (2020) a highly efficient CNN by Google  
098 AI that balances depth, width, and resolution for superior performance and reduced computational  
099 cost in image classification tasks. After that, it was updated utilizing sophisticated DL methods like  
100 the fused mobile inverted bottleneck (Fused-MBConv) operation to produce even better performance  
101 and was named EfficientNetV2 Tan & Le (2021). While ResNet-50 CNN model has been granted  
102 as a promising CNN model for image classification task He et al. (2015b). This study utilized the  
103 EfficientNetV2 architecture and contrasted its performance with ResNet-50 for the classification of  
104 four different stages of kidney cancer image dataset “CT KIDNEY DATASET” Kid (2022). This  
105 study subdivided into four binary classifications: Cyst\_vs\_Normal, Cyst\_vs\_Stone, Cyst\_vs\_Tumor,  
106 and Stone\_vs\_Tumor, aiming to distinguish between specific kidney conditions and streamline the  
107 diagnostic process. Each of them trained to differentiate between them, allowing for more focused  
and precise detection of kidney abnormalities. Those CNN models demonstrate high accuracy across  
all conditions, highlighting their potential for kidney cancer diagnosis and prediction.

Table 1: Kidney images dataset summary at four different stages

Label/Set	Total Count	Training	Validation	Test
<b>Normal vs Cyst</b>				
Normal	5077	4077	479	521
Cyst	3709	2951	399	359
<i>Combined</i>	8786	7028	878	880
<b>Cyst vs Stone</b>				
Cyst	3709	2992	349	368
Stone	1377	1077	158	142
<i>Combined</i>	5086	4069	507	510
<b>Cyst vs Tumor</b>				
Cyst	3709	2993	359	357
Tumor	2283	1800	240	243
<i>Combined</i>	5992	4793	599	600
<b>Stone vs Tumor</b>				
Stone	1377	1084	152	141
Tumor	2283	1844	214	225
<i>Combined</i>	3660	2928	366	366

### 3 RESULT AND DISCUSSION

#### 3.1 DATASET OVERVIEW

The uploaded dataset in Kaggle Kid (2022), primarily sourced from picture archiving and communication system (PACS) at Dhaka, Bangladesh hospital Sharma & Lalwani (2024). Contains both axial cuts and coronal from disparity and non-disparity studies, following protocols for the entire abdomen and urogram, which were selected for collecting the images. Figure 1 displays sample images of kidneys, with red marks highlighting the regions of interest that radiologists use to make specific diagnoses.

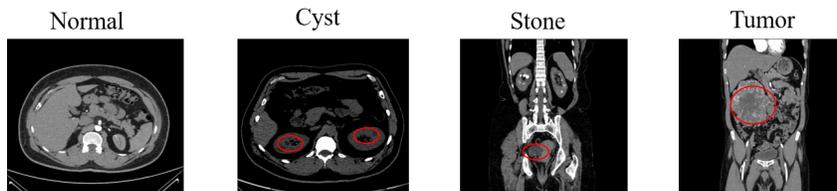


Figure 1: Sample CT scans kidney cancer image data Kid (2022). (a) A normal kidney with a consistent and uniform structure without any abnormalities, serves as a baseline for comparison with the other conditions. (b) Cysts, types of sacs filled with fluid, may vary in size and can sometimes cause pain or other complications. (c) Stone, a hard deposit made of minerals and salts, can cause severe pain and may require medical intervention. (d) A tumor, an abnormal growth of tissue, can be benign or malignant (cancerous), such as RCC. The highlighted areas in red indicate the presence of abnormal growth of tissue

Table 1 presents a dataset summary of the four conditions, detailing the total number of images, in addition to the number of images for training, testing, and validation respectively. The “Normal\_vs\_Cyst” dataset at the top contains the largest number of images, totaling 8786 for training, validation, and testing purposes. The “Cyst\_vs\_Stone” dataset in the middle comprises 5086 images, “Cyst\_vs\_Tumor” dataset below it includes 5992 images. Lastly, the “Stone\_vs\_Tumor” dataset at the bottom consists of 3360 images.

162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215

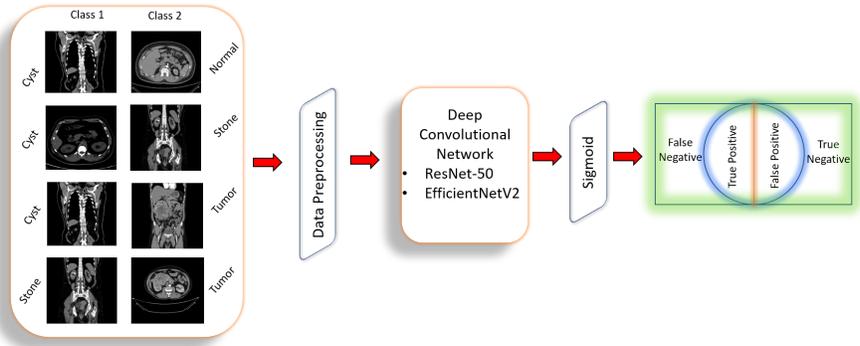


Figure 2: Complete Workflow of an experiment for diagnosing kidney cancer at different conditions. Images are resized to 224×224 as a height and width. After preprocessing the data, deep CNNs He et al. (2015b), Tan & Le (2021) are then developed and trained, with weights optimized through gradient descent. Sigmoid layer outputs provide the score for classifying each image. Finally, the accuracy metric is used to evaluate the model’s performance.

### 3.2 INITIALIZATION OF WEIGHTS

Weight initialization is a critical factor influencing the efficiency and effectiveness of NN training Vedanshu & Tripathi (2018). When weights are initialized to zero, a symmetry problem arises during gradient descent, causing neurons within a layer to learn identical features and hindering effective learning. To address this, specialized initialization methods like Xavier Glorot & Bengio (2010) and He He et al. (2015c) initialization are utilized, each tailored to different activation functions. Xavier initialization, also known as Glorot initialization Glorot & Bengio (2010), is designed for layers using activation functions such as tanh and sigmoid. This method initializes weights using a uniform distribution:

$$W_{Xav} = U \left( -\frac{1}{\sqrt{m}}, \frac{1}{\sqrt{m}} \right) \tag{1}$$

Where the number of input neurons is represented by  $m$ . Xavier initialization helps maintain the variance of activations and gradients throughout the network layers Ramachandran et al. (2017). This promotes smoother convergence during gradient descent, preventing the gradients from vanishing or exploding, which can be crucial for deep networks. For activation functions like ReLU or SiLU Agarap (2019), Xavier initialization can lead to vanishing gradients, particularly for deep networks. He initialization, also known as Kaiming initialization, addresses this by using a Gaussian distribution for weight initialization:

$$W_{He} = \mathcal{N} \left( 0, \sqrt{\frac{2}{m}} \right) \tag{2}$$

where  $\mathcal{N}(0, \sigma^2)$  denotes a normal distribution with mean 0 and  $\sigma^2$  He et al. (2015c). This approach scales the weights to ensure that the variance of the activations remains consistent, which is particularly important for ReLU and similar activation functions that can otherwise suffer from sparse gradients. Modern NN architectures such as ResNet-50 and EfficientNetV2 predominantly use ReLU or SiLU activation functions, as a result, creators of those models used He initialization in these layers He et al. (2015b). Consequently, these models typically employ He initialization to ensure effective training. ResNet-50, leverages residual connections that help mitigate the vanishing gradient problem, and He initialization further complements this by maintaining appropriate gradient scales. EfficientNetV2, emphasizing optimization accuracy and computational efficiency, similarly benefits from He initialization to achieve robust performance.

TL Zhuang et al. (2020) is a powerful technique that significantly enhances training efficiency, especially when data is limited. It involves taking a pre-trained model, often trained on a large dataset like ImageNet, and fine-tuning it for a different but related task. By initializing the model with pre-trained weights, such as those from ImageNet Abadi et al. (2016b), TL allows the model to

leverage previously learned features and patterns. This provides a strong starting point and leads to more effective training with less data Abadi et al. (2016b). For example, when fine-tuning ResNet-50 or EfficientNetV2 for a new task, initializing with pre-trained ImageNet weights enables faster convergence and often results in better performance than training from scratch.

### 3.3 TRAINING RESULT AND DL MODELS

Two modern CNN architectures ResNet-50 and EfficientNetV2 are used, and both models were trained and evaluated using TensorFlow Abadi et al. (2016b); Tan & Le (2021); He et al. (2016). To ensure optimal performance, a TensorFlow checkpoint mechanism was implemented for tracking the accuracy of each epoch during the validation process, which allowed us to keep into record of the topmost model weights after the training process Abadi et al. (2016a). The training for both ResNet-50 and EfficientNetV2 was conducted over 10 epochs with 256 iterations per epoch. The best weights for the ResNet-50 model were saved at epochs 06, 10, 05, and 08 for the Cyst\_vs\_Normal, Cyst\_vs\_Stone, Cyst\_vs\_Tumor, and Stone\_vs\_Tumor datasets, respectively. In contrast, the EfficientNetV2 model achieved its best weights at epochs 09, 10, 07, and 10 for the same datasets.

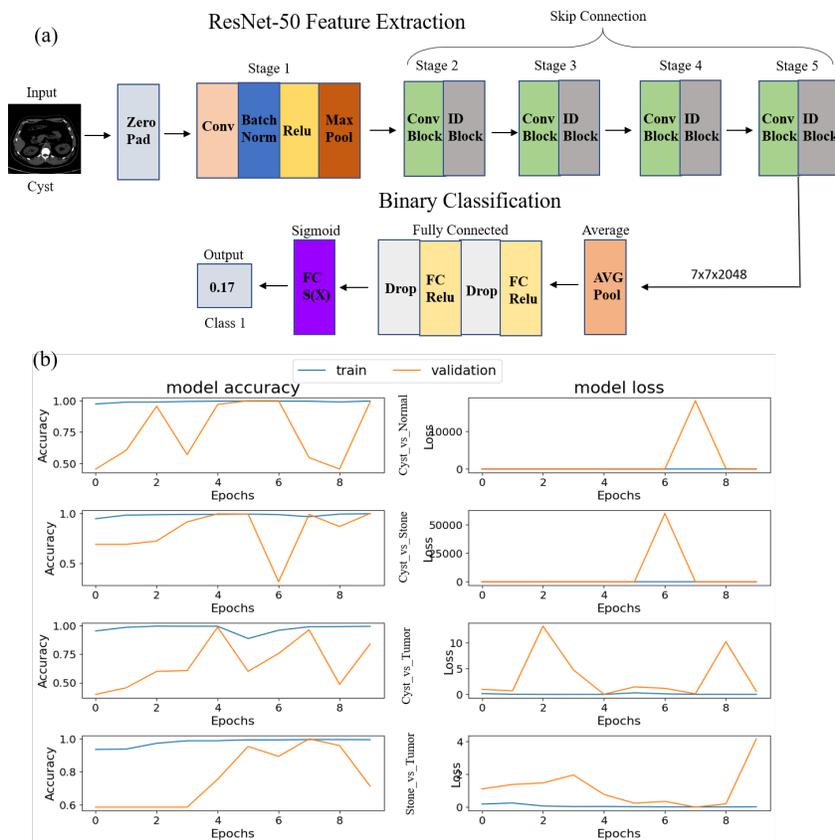


Figure 3: (a) A brief architecture of the Resnet-50 model representing through step-by-step process. The framework demonstrates a forward-passing process through a feature extraction (at the top) and a binary classification stage (at the bottom). (b) Graphical representation of training for four conditions. Left and right column represents the model accuracy and loss after the training (blue line) and validation (orange line) process.

The ResNet-50 model, developed by Microsoft in 2015, is a well-known CNN architecture demonstrating outstanding accomplishment on the ImageNet dataset He et al. (2015b). As shown in Figure 3 (a), the architecture comprises a feature extraction and a binary classification. At the top five distinct stages are included for feature extraction, which recognize the most valuable features from the given input image tensor, Apart from that binary classification layers are employed at the bottom

using the sigmoid activation function  $S(x)$ . During feature extraction, the last four layers (2–5) contain a combination of convolutional and identity blocks He et al. (2015b), Simonyan & Zisserman (2015), as depicted in Figure 3 (a). Each of these blocks incorporates a skip connection, in addition to convolutional layers and batch normalization Ioffe & Szegedy (2015). These skip connections help mitigate issues related to vanishing and exploding gradients, allowing the model to effectively utilize a greater number of layers and learn more complex features, ultimately leading to improved accuracy. After extracting high-level features, the intermediate output image tensor got flattened before passing through the final fully connected layer as illustrated in Figure 3 (a). The feature extractor’s output (originally  $7 \times 7 \times 2048$ ) is flattened out into a one-dimensional array of size 2048, by averaging the first two dimensions using global average pooling Lin et al. (2014). FC and dropout layers then reduce the 2048-node array to a one-dimensional final output of 512 nodes. The final prediction is obtained by the FC layer at the end, which includes a sigmoid activation function.

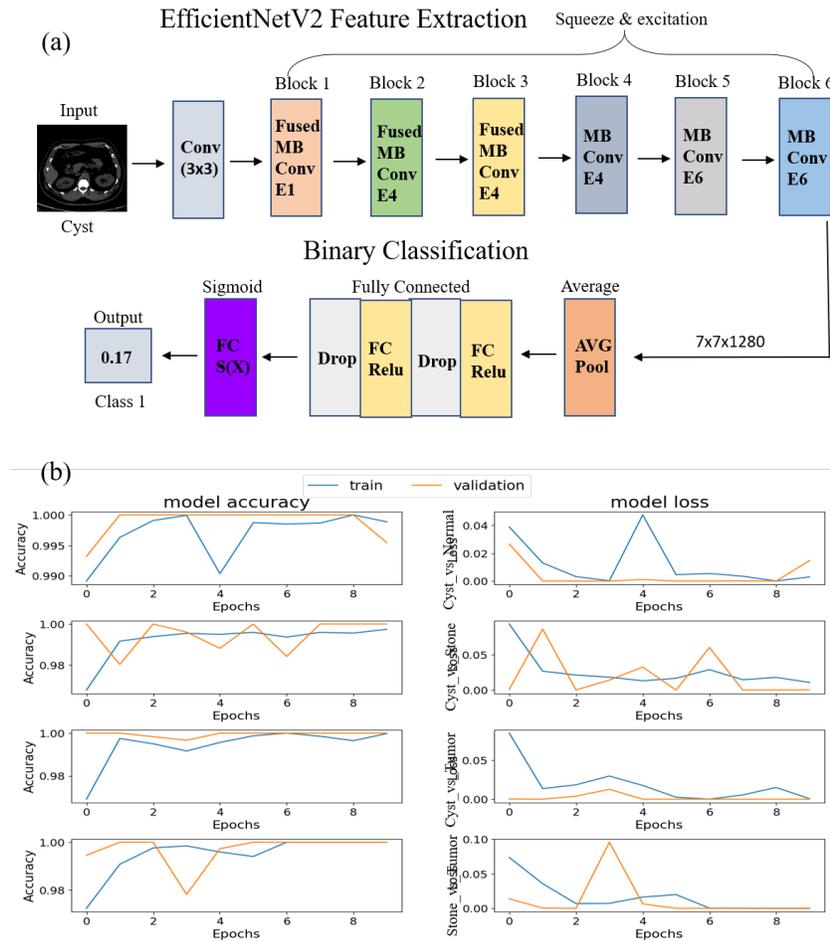


Figure 4: (a) A brief architecture of the EfficientNetV2 model representing through step-by-step process. The framework demonstrates a forward-passing process through a feature extraction stage (at the top) and a binary classification stage (at the bottom). The number of channels at each MB-Conv block is expanded, represented by the expansion component (E). (b) Graphical representation of training for four categories. Left and right column represents the model accuracy and loss after the training (blue line) and validation (orange line) process.

The training and validation process results, shown in Figure 3 (b), demonstrate the ResNet-50 model accuracy and loss for each category over the epochs. The training accuracy and loss consistently trend upward and downward from the start, while validation shows some fluctuations. Implementing a TensorFlow checkpoint Abadi et al. (2016b), the most effective epoch is identified throughout the training process. As a result, the best weights were retained corresponding to the best validation

324 accuracy, ensuring that the validation images were not included in the training phase. This approach  
325 led to improved accuracy in testing.  
326

327 Now for the second model, EfficientNetV2 represents cutting-edge supervised tasks, developed by  
328 Google Brain in 2021 Tan & Le (2021). This model leverages the method, known as neural ar-  
329 chitecture search (NAS) Tan et al. (2019), which systematically explores various ML architectures  
330 to identify the most effective design by sampling different configurations and evaluating their per-  
331 formance. EfficientNetV2 builds upon its predecessor, EfficientNetV1, by optimizing speed during  
332 training and incorporating a valuable operation, Fused-MBConv at the earlier layers. Unlike the  
333 original EfficientNetV1, which employs depthwise convolutions, the Fused-MBConv layer utilizes  
334 standard  $3 \times 3$  convolutions, as illustrated in Figure 4 (a). Additionally, EfficientNetV2 incorporated  
335 strategies from several prior studies to enhance training efficiency keeping the number of parameters  
336 at manageable levels. One such technique is progressively adaptive regularization by learning, pro-  
337 viding a regularized framework that aligns with the image resolution and image size, starting with lower val-  
338 ues in the early epochs and increasing them in later stages. Although progressive learning was not  
339 implemented here, the TL methods reduced the number of parameters and allowed for a significant  
340 reduction in training time approximately less than the ResNet-50 model. The principal components  
341 of the EfficientNetV2 architecture include MBConv and Fused-MBConv layers, which are reused  
342 multiple times throughout the model. As depicted in Figure 4 (a), the architecture is divided into  
343 segments for feature extraction and binary classification similar to ResNet-50. The feature extractor  
344 begins with a stem that includes a standard convolution layer, followed by six distinct blocks that  
345 consist of several repetitions of Fused-MBConv and MBConv layers. After collecting feature tensors  
346 of size  $7 \times 7 \times 1280$ , global average pooling is applied Lin et al. (2014), resulting in a single value  
347 per channel. Then passing through three layers (two ReLU activation and one sigmoid activation)  
output is obtained.

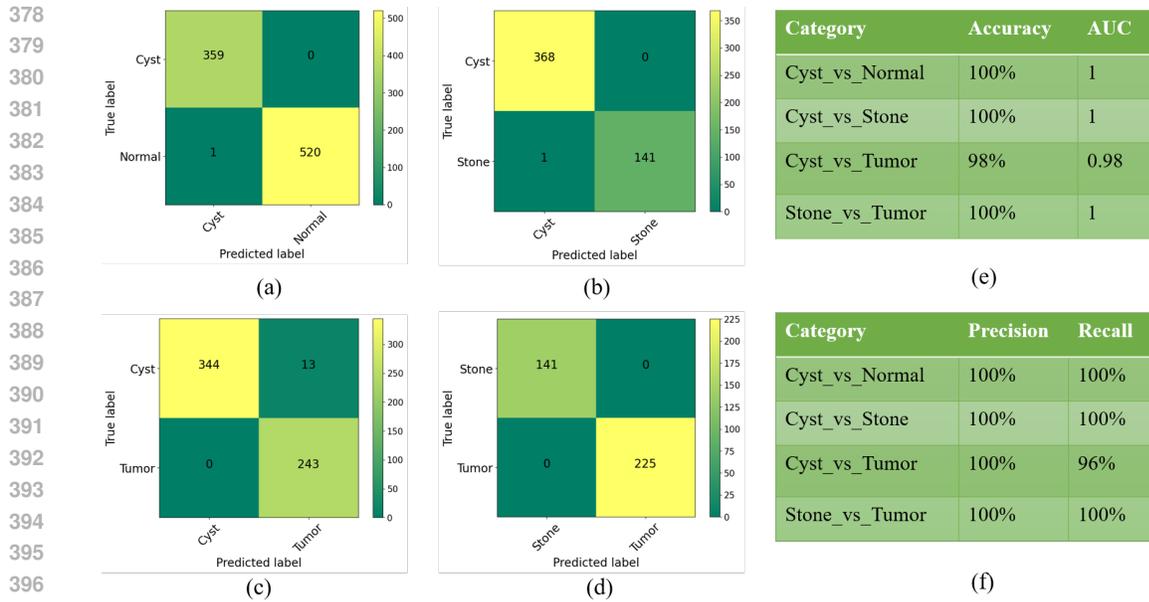
348 The training and validation process results, shown in Figure 4 (b) demonstrate the model accuracy  
349 and loss over each category across the epochs. The Cyst\_vs\_Normal dataset is shown at the top,  
350 followed by Cyst\_vs.Stone and Cyst\_vs.Tumor in the middle, and Stone\_vs.Tumor at the bottom.  
351 Similar to ResNet-50 the most effective epoch is identified throughout the training process, by im-  
352 plementing a TensorFlow checkpoint Abadi et al. (2016b). As a result, the best weights are retained  
353 corresponding to the best validation accuracy. While the training time of EfficientNetV2 is much  
354 faster than the ResNet-50. Overall, the EfficientNetV2 model demonstrates comparable stability in  
355 accuracy and loss to the ResNet-50 model. Training with EfficientNetV2 was approximately faster  
356 than ResNet-50, thanks to optimizations introduced through neural architecture search (NAS).  
357

### 358 3.4 MODEL PERFORMANCE 359

360 The performance of the ResNet-50 model across four binary classifications demonstrates its high  
361 reliability in medical image classification [see Figure 5]. The model achieves perfect or near-perfect  
362 accuracy in all four conditions, with minimal misclassifications, highlighting its robust capability in  
363 distinguishing between the conditions. The Cyst\_vs.Tumor task shows a slight decrease in perfor-  
364 mance compared to the others, reflecting the inherent challenge in differentiating these two classes.  
365 Precision remains consistently at 100% for all tasks, indicating the absence of false positives, while  
366 recall is also perfect except for a minor reduction in the Cyst\_vs.Tumor task. These results estab-  
367 lish ResNet-50 as a highly effective tool for binary classification in medical imaging, capable of  
368 delivering reliable predictions with minimal errors across diverse categories.

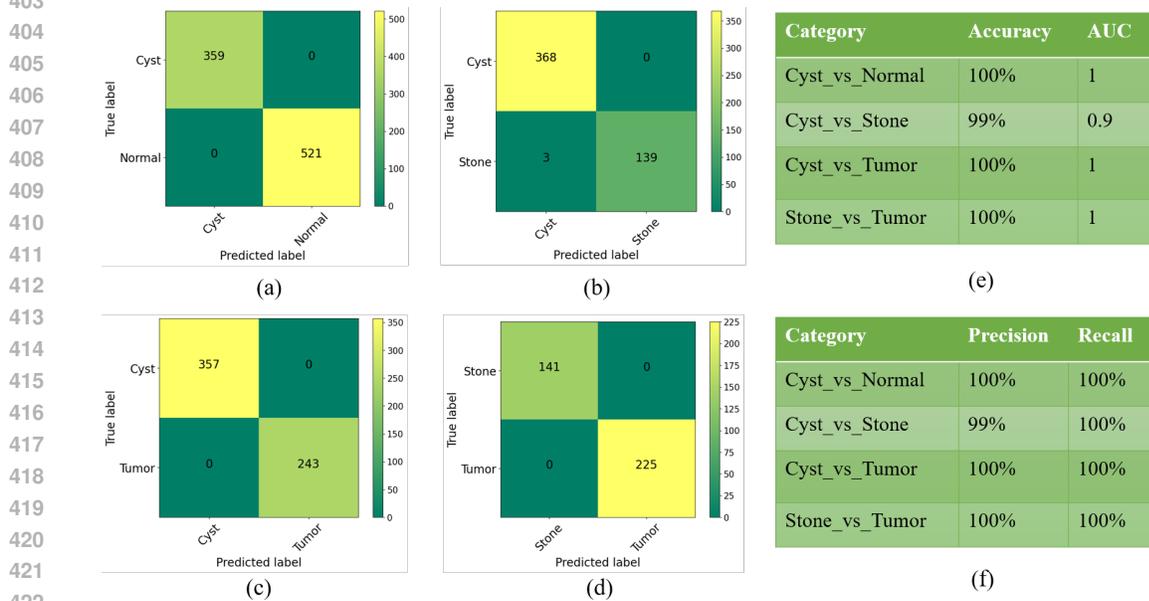
369 In contrast, the EfficientNetV2 model demonstrates strong performance across the four binary clas-  
370 sification tasks [see Figure 6]. It achieves perfect accuracy for all tasks and near-perfect accuracy in  
371 the Cyst\_vs.Stone. Precision remains consistently high across all of the classifications, with only a  
372 slight drop in the Cyst\_vs.Stone. Similarly, recall is perfect across all classifications. These results  
373 underscore the model’s reliability and effectiveness in medical image classification, with minimal  
374 errors, making it well-suited for clinical applications.

375 Both models exhibit similarly high performance across all four binary classification tasks, but Ef-  
376 ficientNetV2 shows a slight advantage in the Cyst\_vs.Tumor, where it achieves flawless classifica-  
377 tion, unlike ResNet-50, which had minor misclassifications. However, EfficientNetV2 experienced a  
small dip in precision during the Cyst\_vs.Stone task, where ResNet-50 maintained perfect precision.



398  
399  
400  
401  
402  
403

Figure 5: Performance of ResNet-50 model over the test sets of four different conditions. (a) Cyst\_vs\_Normal (b) Cyst\_vs\_Stone (c) Cyst\_vs\_Tumor (d) Stone\_vs\_Tumor represents binary confusion matrix. (e) shows the accuracy and AUC score for each binary classification, while (f) displays the corresponding precision and recall metrics.



424  
425  
426  
427  
428  
429

Figure 6: Performance of EfficientNetV2 model over the test set. (a) Cyst\_vs\_Normal (b) Cyst\_vs\_Stone (c) Cyst\_vs\_Tumor (d) Stone\_vs\_Tumor represents binary confusion matrix. (e) shows the accuracy AUC score for each binary classification, while panel (f) displays the corresponding precision and recall metrics.

430  
431

Overall, both models deliver remarkable accuracy, EfficientNetV2 shows a slightly more consistent recall performance across all tasks, making it a marginally more robust model for medical image classification, especially when subtle class distinctions are involved. Even small gains in accuracy,

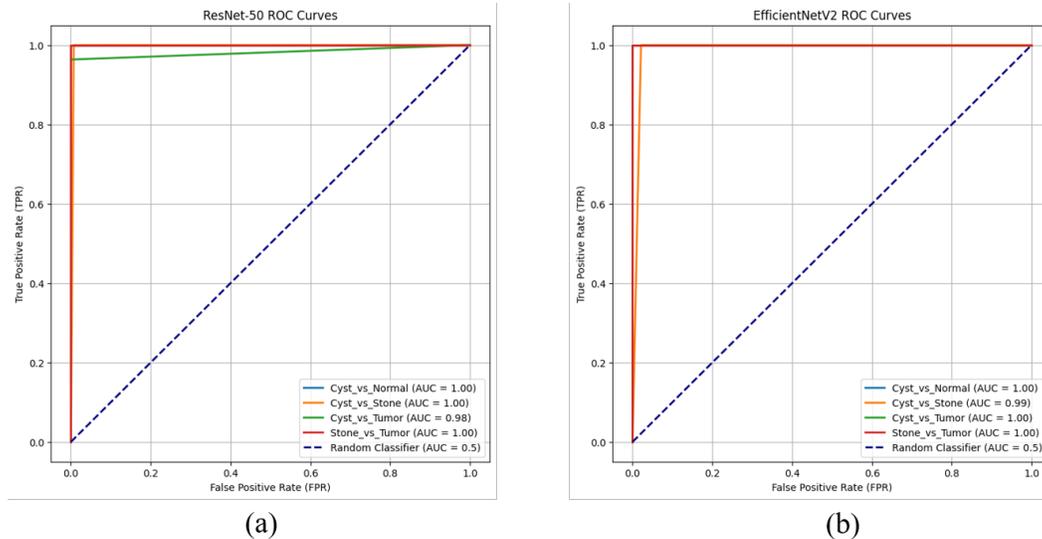


Figure 7: Receiver Operating Characteristic (ROC) curves for two deep CNN models: (a) ResNet-50 and (b) EfficientNetV2, across multiple binary classification tasks. The plots represent the True Positive Rate (TPR) against the False Positive Rate (FPR) for the four different conditions: Cyst\_vs\_Normal, Cyst\_vs\_Stone, Cyst\_vs\_Tumor, and Stone\_vs\_Tumor, along with a reference random classifier (AUC = 0.5). The AUC values indicate excellent classification performance, with most tasks achieving AUC values close to 1.

as seen with models like EfficientNetV2 compared to ResNet-50, could have significant implications for future medical applications. The high levels of accuracy achieved with these datasets highlight the potential of deep CNNs in medical image analysis. These findings point toward a promising research direction in the precise identification of renal disease subtypes, where the ability of CNNs to detect subtle variations in images could lead to more effective and reliable diagnostic tools in healthcare.

## 4 CONCLUSION

The DL frameworks presented in this study effectively address the critical need for early detection of renal diseases. By leveraging TL and advanced CNNs system achieved remarkable testing accuracy of up to 100% across multiple classifications. These findings not only demonstrate the potential for precise and automatic classification of renal conditions but also highlight the framework’s applicability in clinical settings. By improving early diagnosis, this approach could lead to better patient management, reduced rates of chronic kidney disease, and ultimately improved survival rates in renal cancer. However, future work is needed to inflate the robustness and generalizability of the model. This includes expanding the dataset to encompass a wider variety of demographic and clinical variables, integrating multimodal imaging techniques, and conducting prospective clinical trials to validate performance in real-world scenarios. Additionally, further research could focus on the interpretability of the model’s predictions to support clinicians in decision-making processes.

## REFERENCES

<https://gco.iarc.who.int/media/globocan/factsheets/cancers/29-kidney-fact-sheet.pdf>, 2022.

<https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone>, 2022.

Marín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg,

- 486 Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan,  
487 Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. Tensorflow: A system for large-  
488 scale machine learning, 2016a. URL <https://arxiv.org/abs/1605.08695>.  
489
- 490 Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu  
491 Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg,  
492 Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasude-  
493 van, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: A system for  
494 large-scale machine learning, May 2016b. URL <http://arxiv.org/abs/1605.08695>.  
495 arXiv:1605.08695 [cs].
- 496 Abien Fred Agarap. Deep Learning using Rectified Linear Units (ReLU), February 2019. URL  
497 <http://arxiv.org/abs/1803.08375>. arXiv:1803.08375 [cs, stat].
- 498 Işıl Aksakalli, Sibel Kaçdioğlu, and Y. Sinan Hanay. Kidney X-ray Images Classification using  
499 Machine Learning and Deep Learning Methods. *Balkan Journal of Electrical and Computer*  
500 *Engineering*, 9(2):144–151, April 2021. ISSN 2147-284X. doi: 10.17694/bajece.878116. URL  
501 <http://dergipark.org.tr/en/doi/10.17694/bajece.878116>.
- 502 Tilahun Alelign and Beyene Petros. Kidney Stone Disease: An Update on Current Concepts. *Ad-*  
503 *vances in Urology*, 2018:1–12, 2018. ISSN 1687-6369, 1687-6377. doi: 10.1155/2018/3068365.  
504 URL <https://www.hindawi.com/journals/au/2018/3068365/>.  
505
- 506 Yonghua Bi, Xiaonan Shi, Jianzhuang Ren, Mengfei Yi, and Xinwei Han. Transarterial chemoem-  
507 bolization of unresectable renal cell carcinoma with doxorubicin-loaded CalliSpheres drug-  
508 eluting beads. *Scientific Reports*, 12(1):8136, May 2022. ISSN 2045-2322. doi: 10.1038/  
509 s41598-022-12334-x. URL <https://doi.org/10.1038/s41598-022-12334-x>.
- 510 Neta Blau, Eyal Klang, Nahum Kiryati, Marianne Amitai, Orith Portnoy, and Arnaldo Mayer.  
511 Fully automatic detection of renal cysts in abdominal CT scans. *International Journal of Com-*  
512 *puter Assisted Radiology and Surgery*, 13(7):957–966, July 2018. ISSN 1861-6410, 1861-  
513 6429. doi: 10.1007/s11548-018-1726-6. URL [http://link.springer.com/10.1007/  
514 s11548-018-1726-6](http://link.springer.com/10.1007/s11548-018-1726-6).
- 515 Kyle J Foreman, Neal Marquez, Andrew Dolgert, Kai Fukutaki, Nancy Fullman, Madeline Mc-  
516 Gaughey, Martin A Pletcher, Amanda E Smith, Kendrick Tang, Chun-Wei Yuan, Jonathan C  
517 Brown, Joseph Friedman, Jiawei He, Kyle R Heuton, Mollie Holmberg, Disha J Patel, Patrick  
518 Reidy, Austin Carter, Kelly Cercy, Abigail Chapin, Dirk Douwes-Schultz, Tahvi Frank, Falko  
519 Goettsch, Patrick Y Liu, Vishnu Nandakumar, Marissa B Reitsma, Vince Reuter, Nafis Sa-  
520 dat, Reed J D Sorensen, Vinay Srinivasan, Rachel L Updike, Hunter York, Alan D Lopez,  
521 Rafael Lozano, Stephen S Lim, Ali H Mokdad, Stein Emil Vollset, and Christopher J L Mur-  
522 ray. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortal-  
523 ity for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 coun-  
524 tries and territories. *The Lancet*, 392(10159):2052–2090, November 2018. ISSN 01406736.  
525 doi: 10.1016/S0140-6736(18)31694-5. URL [https://linkinghub.elsevier.com/  
526 retrieve/pii/S0140673618316945](https://linkinghub.elsevier.com/retrieve/pii/S0140673618316945).
- 527 Xu Fu, Huaiqin Liu, Xiaowang Bi, and Xiao Gong. Deep-Learning-Based CT Imaging in the Quan-  
528 titative Evaluation of Chronic Kidney Diseases. *Journal of Healthcare Engineering*, 2021:1–  
529 9, October 2021. ISSN 2040-2309, 2040-2295. doi: 10.1155/2021/3774423. URL [https://  
530 www.hindawi.com/journals/jhe/2021/3774423/](https://www.hindawi.com/journals/jhe/2021/3774423/).
- 531 Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural  
532 networks. In Yee Whye Teh and Mike Titterton (eds.), *Proceedings of the Thirteenth Interna-*  
533 *tional Conference on Artificial Intelligence and Statistics*, volume 9 of *Proceedings of Machine*  
534 *Learning Research*, pp. 249–256, Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010. PMLR.  
535 URL <https://proceedings.mlr.press/v9/glorot10a.html>.
- 536 Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recog-  
537 nition, 2015a. URL <https://arxiv.org/abs/1512.03385>.
- 538 Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image  
539 Recognition, 2015b. URL <https://arxiv.org/abs/1512.03385>. Version Number: 1.

- 540 Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving Deep into Rectifiers: Sur-  
541 passing Human-Level Performance on ImageNet Classification, February 2015c. URL [http://](http://arxiv.org/abs/1502.01852)  
542 [arxiv.org/abs/1502.01852](http://arxiv.org/abs/1502.01852). arXiv:1502.01852 [cs].  
543
- 544 Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recog-  
545 nition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*  
546 (*CVPR*), June 2016.
- 547 Mohammad Sakib Hossain, S. M. Nazmul Hassan, Mohammad Al-Amin, Md. Nakib Rahaman,  
548 Rakib Hossain, and Muhammad Iqbal Hossain. Kidney disease detection from ct images using  
549 a customized cnn model and deep learning. In *2023 International Conference on Advances in*  
550 *Intelligent Computing and Applications (AICAPS)*, pp. 1–6, 2023. doi: 10.1109/AICAPS57044.  
551 2023.10074314.
- 552 James J. Hsieh, Mark P. Purdue, Sabina Signoretti, Charles Swanton, Laurence Albiges, Manuela  
553 Schmidinger, Daniel Y. Heng, James Larkin, and Vincenzo Ficarra. Renal cell carcinoma. *Nature*  
554 *Reviews Disease Primers*, 3(1):17009, March 2017. ISSN 2056-676X. doi: 10.1038/nrdp.2017.9.  
555 URL <https://www.nature.com/articles/nrdp20179>.  
556
- 557 Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by  
558 reducing internal covariate shift, 2015. URL <https://arxiv.org/abs/1502.03167>.
- 559 Stefan Jacobson. [Chronic kidney disease—a public health problem]. *Lakartidningen*, 110(21):1018–  
560 1020, May 2013. ISSN 0023-7205.
- 561 Asifullah Khan, Anabia Sohail, Umme Zahoor, and Aqsa Saeed Qureshi. A survey of the recent  
562 architectures of deep convolutional neural networks. *Artificial Intelligence Review*, 53(8):5455–  
563 5516, December 2020. ISSN 0269-2821, 1573-7462. doi: 10.1007/s10462-020-09825-6. URL  
564 <https://link.springer.com/10.1007/s10462-020-09825-6>.  
565
- 566 Min Lin, Qiang Chen, and Shuicheng Yan. Network In Network, March 2014. URL [http://](http://arxiv.org/abs/1312.4400)  
567 [arxiv.org/abs/1312.4400](http://arxiv.org/abs/1312.4400). arXiv:1312.4400 [cs].  
568
- 569 Vishal Navani and Daniel Y C Heng. Immunotherapy in renal cell carcinoma. *The*  
570 *Lancet Oncology*, 24(11):1164–1166, November 2023. ISSN 14702045. doi: 10.1016/  
571 S1470-2045(23)00473-4. URL [https://linkinghub.elsevier.com/retrieve/](https://linkinghub.elsevier.com/retrieve/pii/S1470204523004734)  
572 [pii/S1470204523004734](https://linkinghub.elsevier.com/retrieve/pii/S1470204523004734).
- 573 Anushri Parakh, Hyunkwang Lee, Jeong Hyun Lee, Brian H. Eisner, Dushyant V. Sahani, and  
574 Synho Do. Urinary Stone Detection on CT Images Using Deep Convolutional Neural Net-  
575 works: Evaluation of Model Performance and Generalization. *Radiology: Artificial Intelli-*  
576 *gence*, 1(4):e180066, July 2019. ISSN 2638-6100. doi: 10.1148/ryai.2019180066. URL  
577 <http://pubs.rsna.org/doi/10.1148/ryai.2019180066>.  
578
- 579 Mallikarjun Raghavendra and M Vidya. Functions of kidney & artificial kidneys. *International*  
580 *Journal of Innovative Research*, 1(11):1–5, 2013.
- 581 Prajit Ramachandran, Barret Zoph, and Quoc V. Le. Searching for Activation Functions, October  
582 2017. URL <http://arxiv.org/abs/1710.05941>. arXiv:1710.05941 [cs].  
583
- 584 Stuart S. Sagel, Robert J. Stanley, Robert G. Levitt, and Guillermo Geisse. Computed Tomography  
585 of the Kidney. *Radiology*, 124(2):359–370, August 1977. ISSN 0033-8419, 1527-1315. doi:  
586 10.1148/124.2.359. URL <http://pubs.rsna.org/doi/10.1148/124.2.359>.
- 587 Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mo-  
588 bilenetv2: Inverted residuals and linear bottlenecks, 2019. URL [https://arxiv.org/abs/](https://arxiv.org/abs/1801.04381)  
589 [1801.04381](https://arxiv.org/abs/1801.04381).
- 590 Neeraj Sharma and Praveen Lalwani. Predictor model for chronic kidney disease using adaptive  
591 gradient clipping with deep neural nets. *International Journal of Advanced Computer Science*  
592 *and Applications*, 15(4), 2024. doi: 10.14569/IJACSA.2024.01504119. URL [http://dx.](http://dx.doi.org/10.14569/IJACSA.2024.01504119)  
593 [doi.org/10.14569/IJACSA.2024.01504119](http://dx.doi.org/10.14569/IJACSA.2024.01504119).

- 594 Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image  
595 recognition, 2015. URL <https://arxiv.org/abs/1409.1556>.  
596
- 597 Stephen M. Sozio, Kurtis A. Pivert, Fergus J. Caskey, and Adeera Levin. The state of the global  
598 nephrology workforce: a joint ASN-ERA-EDTA-ISN investigation. *Kidney International*, 100  
599 (5):995–1000, November 2021. ISSN 00852538. doi: 10.1016/j.kint.2021.07.029. URL <https://linkinghub.elsevier.com/retrieve/pii/S0085253821007638>.  
600
- 601 S Sudharson and Priyanka Kokil. An ensemble of deep neural networks for kidney ultrasound  
602 image classification. *Computer Methods and Programs in Biomedicine*, 197:105709, December  
603 2020. ISSN 01692607. doi: 10.1016/j.cmpb.2020.105709. URL <https://linkinghub.elsevier.com/retrieve/pii/S016926072031542X>.  
604
- 605 Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Du-  
606 mitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions,  
607 2014. URL <https://arxiv.org/abs/1409.4842>.  
608
- 609 Mingxing Tan and Quoc V. Le. Efficientnet: Rethinking model scaling for convolutional neural  
610 networks, 2020. URL <https://arxiv.org/abs/1905.11946>.  
611
- 612 Mingxing Tan and Quoc V. Le. EfficientNetV2: Smaller Models and Faster Training, June 2021.  
613 URL <http://arxiv.org/abs/2104.00298>. arXiv:2104.00298 [cs].
- 614 Mingxing Tan, Bo Chen, Ruoming Pang, Vijay Vasudevan, Mark Sandler, Andrew Howard, and  
615 Quoc V. Le. Mnasnet: Platform-aware neural architecture search for mobile. In *2019 IEEE/CVF*  
616 *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2815–2823, 2019. doi:  
617 10.1109/CVPR.2019.00293.
- 618 Vedanshu and M. M. Tripathi. Zero Initialization of modified Gated Recurrent Encoder-Decoder  
619 Network for Short Term Load Forecasting, December 2018. URL <http://arxiv.org/abs/1812.03425>. arXiv:1812.03425 [cs, stat].  
620
- 621 Jyoti Verma, Madhwendra Nath, Priyanshu Tripathi, and K. K. Saini. Analysis and identification of  
622 kidney stone using Kth nearest neighbour (KNN) and support vector machine (SVM) classifica-  
623 tion techniques. *Pattern Recognition and Image Analysis*, 27(3):574–580, July 2017. ISSN 1054-  
624 6618, 1555-6212. doi: 10.1134/S1054661817030294. URL <http://link.springer.com/10.1134/S1054661817030294>.  
625
- 626 Shen Wang, Yuyuan Zhou, Xiaochen Qin, Suresh Nair, Xiaolei Huang, and Yaling Liu. Label-  
627 free detection of rare circulating tumor cells by image analysis and machine learning. *Scientific*  
628 *Reports*, 10(1):12226, July 2020. ISSN 2045-2322. doi: 10.1038/s41598-020-69056-1. URL  
629 <https://doi.org/10.1038/s41598-020-69056-1>.  
630
- 631 Kadir Yildirim, Pinar Gundogan Bozdag, Muhammed Talo, Ozal Yildirim, Murat Karabatak, and  
632 U.Rajendra Acharya. Deep learning model for automated kidney stone detection using coronal  
633 CT images. *Computers in Biology and Medicine*, 135:104569, August 2021. ISSN 00104825.  
634 doi: 10.1016/j.compbiomed.2021.104569. URL <https://linkinghub.elsevier.com/retrieve/pii/S0010482521003632>.  
635
- 636 Q. Zheng, S.L. Furth, G.E. Tasian, and Y. Fan. Computer-aided diagnosis of congenital abnormal-  
637 ities of the kidney and urinary tract in children based on ultrasound imaging data by integrating  
638 texture image features and deep transfer learning image features. *Journal of Pediatric Urology*,  
639 15(1):75.e1–75.e7, February 2019. ISSN 14775131. doi: 10.1016/j.jpuro.2018.10.020. URL  
640 <https://linkinghub.elsevier.com/retrieve/pii/S1477513118303541>.  
641
- 642 Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and  
643 Qing He. A Comprehensive Survey on Transfer Learning, June 2020. URL <http://arxiv.org/abs/1911.02685>. arXiv:1911.02685 [cs, stat].  
644
- 645
- 646
- 647