Journal of Al-Turath University College (JATUC)
Open Access Journal
ISSN : 2074-5621
https://jatuc.uoturath.edu.iq/

Hybrid K-means and CNN Refinement for Kidney Images Segmentation

Bushra Raad Zahi[*]

Department of Information Technology Management, Technical College of Management /Baghdad, Middle Technical University, Baghdad, Iraq.

[*]*Correspondence author*: Bushra_raad@mtu.edu.iq

KEYWORDS

Image segmentation, K-means clustering, CNN refinement, hybrid methods, real-time processing.

ABSTRACT

This study aims to present a two-stage framework combining artificial intelligence algorithms, including the efficient K-means algorithm and a lightweight convolutional neural network (CNN) algorithm, which provides high accuracy for precise segmentation of Ultrasound images and obtaining more accurate data. Image segmentation is a fundamental process for image analysis and obtaining more precise information; it is one of the best methods used in image processing. The K-means algorithm generates an initial mask (an initial image segmentation). The CNN algorithm, a neural network, then refines this mask to recover the precise boundaries of objects within the overall image, thus identifying the image's edges without loss of features and pinpointing the location for diagnosis. The results obtained from the Ultrasound images used on the BSD500 and PASCAL VOC 2012 datasets showed a clear improvement compared to using K-means alone, as the CNN algorithm proved its worth in obtaining more accurate and clear data, and achieving performance that rivals the latest deep learning-based methods, while maintaining real-time execution speed at the same time.

الكلمات المفتاحية

تقسيم الصور، تجميع الألوان بطريقة K-means، تحسين أداء الشبكات العصبية التفافية (CNN)، الطرق المدمجة أو الهجينة، المعالجة في الوقت الفعلي.

الملخص

تهدف هذه الدراسة الى بيان اطار عمل يتكون من مرحلتين يجمع بين خوارزميات الذكاء الاصطناعي منها خوارزمية K-means التي توفر الكفاءة وشبكة عصبية التفافية خفيفة وهي خوارزمية (CNN) التي توفر الدقة العالية من أجل تحقيق تجزئة دقيقة لِصور الكلى الماخوذة بجهاز السونار وللحصول على بيانات اكثر دقة.حيث ان عملية تجزئة الصورة هي عملية اساسية لتحليل الصور وللحصول على معلومات ادق هي من افضل الطرق المستخدمة في معالجة الصورة . تقوم خوارزمية K-means بإنتاج قناع أوّلي وهو (تقسيم أولي للصورة)، ثم تتولى بعد ذلك خوارزمية الـ CNNوهي شبكة عصبية لتحسين هذا القناع لاستعادة الحدود الدقيقة للكائنات داخل صور الكلى وذلك لتحديد الحواف الخاصة بالصورة وعدم فقدان الملامح الخاصة بها وتحديد المكان المراد تشخيصه. أظهرت النتائج التي اجريت على صور السونار المستخدمة على مجموعتي البيانات BSD500 و PASCAL VOC 2012تحسّنًا واضحًا مقارنةً باستخدام K-means وحدها حيث ان خوارزمية CNN اثبتت جدارة في الحصول على بيانات اكثر دقة وَ وضوح، و تحقيق أداء منافس لأحدث الطرق المعتمدة على التعلّم العميق، مع الحفاظ على سرعة تنفيذ آنية (Real-time) في الوقت ذاته.

1. INTRODUCTION

Medical image segmentation is essential in quantitative analysis of medical images, which can assist a wide variety of clinical procedures, including diagnosis, radiation therapy planning, and surgery [1, 2]. Several segmentation techniques have been developed in the last few decades, from classical methods to more recent deep learning and hybrid models [1, 3]. Traditional image segmentation methods, such as thresholding, region growing, and clustering, have provided a firm foundation in medical image analysis due to their conceptual simplicity and computational feasibility. However, such methods usually have obvious limitations when dealing with the complexity and uncertainty within medical images, particularly in noise, low-contrast, and blur situation[1, 3].

The images used in this study were obtained from a database (Mendeley Data) containing high-quality, high-resolution images suitable for image processing tasks. This study was applied to this type of image of an important organ in the human body, namely the kidney.

The kidneys are two bean-shaped, retroperitoneal organs located on either side of the vertebral column beneath the rib cage. They play a vital role in filtering blood, removing waste products, and maintaining fluid and electrolyte balance [4].

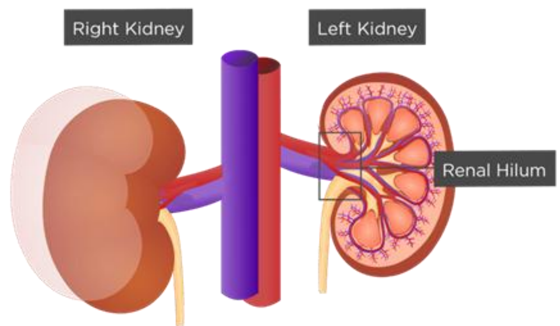

Figure 1: The kidneys in the human body

The arrival of deep learning has produced seminal contributions to the field by enabling the automated extraction of multi-level features from data, and subsequently resulting in remarkable increases in segmentation accuracy in multiple medical imaging modalities [5, 6]. Convolutional Neural Networks (CNNs) and tailored variations for bio-medical applications have set new segmentation benchmarks [6, 7]. However, even with these advances, deep learning models typically need a large annotated data set and can be difficult to interpret, therefore being of continued interest in complementary methods. Fusion-based segmentation approaches have now been proposed, such as 13, that allows to combine the strengths and alleviating the weaknesses of both categories [1, 6, 3, 2]. These models make use of domain expertise, prior knowledge, or operator learning to enhance the robustness, generalizability, and computational effectiveness of the model [6]. For better segmentation results, it is best to combine traditional methods with data-driven ones. This is especially true when working with limited training data and complex structures [7].

The RGB colour space, or more advanced colour spaces like CIELAB, are typically used by colour segmentation methods. Unsupervised methods, such as k-means clustering, are easy to use and perform well when used for this kind of problem. Data with complex textures or hidden objects perform poorly when using K-means since it ignores pixel spatial information. Convolutional neural networks (CNNs) are the state-of-the-art method for image segmentation, according to their basic principle. Highly accurate, complicated segmentation is now possible with the help of CNNs, which can learn spatial and contextual insights on their own. However, CNNs frequently require large labelled datasets and sufficient computational resources. Recently, the modeling of hybrid models for image segmentation has been a major focus in research by integrating classic clustering methods with deep one. It has introduced the combination of superpixel clustering with CNN nothingness for precise boundary detection in other examples [8]. Additional methodologies examine certain loss functions for segmentation, such as boundary-aware or focal loss, to address class imbalance [9].

The U-Net architecture has been a major pillar that can be combined, and widely used for segmentation of natural and

biomedical images due to its successful skip-connection that allows retention of spatial resolution [10]. Moreover, research that improves k-means clustering initialization and optimization [11] result in a more stable clustering output, which greatly benefits hybrid approaches.
In this work, we propose a hybrid approach by integrating clustering and CNNs for feature extraction, as shown in [13]. K-means algorithm is performed for the first segmentation step, and a CNN model for the second. This approach takes advantage of convolutional neural networks' (CNNs') superior segmentation skills with complicated boundaries while making use of clustering's speed and simplicity. The methods of image segmentation can be roughly classified into the following categories:

• Clustering methods: Techniques such as k-means, mean shift, and fuzzy c-means group pixels based on their similarity in color. Fast and unsupervised, these methods, however, neglect texture and edge continuity.

• Deep Learning Methodologies: Fully Convolutional Networks (FCNs) [14,15]. Mask R-CNN are some of the most commonly used deep learning models, which segments by learning features hierarchically from the input image.

• Hybrid based Models: The idea of clustering and learning has drawn attention; the models can use initial region proposals and subsequently refine them by learning. There are a lot of efforts to combine the traditional clustering methods with deep learning.

The rest of this paper is structured as follows. Section 2 presents the proposed segmentation framework, describing k-means clustering and CNN refinement. 3 Datasets and experiment Section 3 introduces the data sets, evaluation protocols, and experimental setup for evaluating the method. The effectiveness of our approach is evidenced by quantitative and qualitative results in Section 4. The paper is concluded in Section 5, which also points to future work.

2. MATERIAL AND METHODS

In this section, we describe the hybrid image segmentation, based on the application of clustering and CNNs. The proposed approach consists of two main steps:
• Clustering, we fit images directed by some colour similarity, based on the k-means algorithm.
• CNN to improve over first-line segmentation: We use a convolutional network that includes the contours of segmentations, complex spatial relationship and outliers.
We want the best of both worlds, which are cheap and weak clustering methods at a computational cost, while CNN can greatly improve their segmentation performance.

2.1. Clustering Stage (K-means)

Problem formulation.
Let the input image be represented by a set of N pixels

$$\mathcal{X} = \{\mathbf{x}_i \in \mathbb{R}^d | i = 1, \dots, N\},$$

where each feature vector $\mathbf{x}_i$ may combine color components (e.g., CIELab or RGB) and optional spatial coordinates (u_i, v_i) to encourage spatially coherent clusters:

$$\mathbf{x}_i = [L_i, a_i, b_i, \lambda\, u_i, \lambda\, v_i]^\top.$$

The scaling factor $\lambda > 0$ balances color and spatial influence.

Given a prescribed number of clusters K, K-means seeks the set of centroids $\boldsymbol{\mu} = \{\boldsymbol{\mu}_k\}_{k=1}^{K}$, $\boldsymbol{\mu}_k \in \mathbb{R}^d$, that minimizes the within-cluster sum of squared errors (WCSS):

$$\mathcal{J}(\boldsymbol{\mu}) = \sum_{k=1}^{K} \sum_{i=1}^{N} r_{ik}\, \|\mathbf{x}_i - \boldsymbol{\mu}_k\|_2^2, \tag{1}$$

where the binary assignment variable

$$r_{ik} = \begin{cases} 1, & \text{if} k = \underset{\ell}{\operatorname{argmin}}\ \| \mathbf{x}_i - \boldsymbol{\mu}_\ell \|_2^2, \\ 0, & \text{otherwise.} \end{cases}$$

Iterative optimization.
Equation (1) is minimized via alternating updates until convergence:

1. *Assignment step*

$$r_{ik} \leftarrow \begin{cases} 1, & k = \operatorname*{argmin}_{\ell} \parallel \mathbf{x}_i - \boldsymbol{\mu}_\ell \parallel_2^2, \\ 0, & \text{otherwise}; \end{cases} \qquad \forall i, k.$$

2. *Centroid update step*

$$\boldsymbol{\mu}_k \;\leftarrow\; \frac{1}{N_k} \sum_{i=1}^{N} r_{ik}\, \mathbf{x}_i, \qquad N_k = \sum_{i=1}^{N} r_{ik}.$$

3. *Stopping criterion* Stop if $\max_k \parallel \boldsymbol{\mu}_k^{(t)} - \boldsymbol{\mu}_k^{(t-1)} \parallel_2 < \varepsilon$ or after $T_{\max}$ iterations.

In practice, we initialized $\boldsymbol{\mu}$ with *K-means++* to improve convergence, and repeated the procedure R times ($R = 5$ in our experiments), retaining the solution with the lowest $\mathcal{J}$.

Segmentation mask generation.

After convergence, each pixel inherits the label of its nearest centroid $k^\star$, producing a hard segmentation map

$$S\colon I \longrightarrow \{1, \ldots, K\}, \qquad S(\mathbf{x}_i) = k^\star.$$

This coarse mask is subsequently refined by the *CNN refinement stage* (see Section 2.2) to capture fine-grained object boundaries.

2.1.1 Color Space Transformation

We start by converting the image from RGB color space to **CIELAB** color space. The CIELAB space is selected because it is perceptually uniform, the distances between points in this space correspond more closely with human perception.

The transformation from RGB to LAB is given by:

$$\mathbf{I}_{LAB} = T_{RGB\to LAB}(\mathbf{I}),$$

where **I** represents the input image, and $T_{RGB\to LAB}$ is the transformation function.

2.1.2 Feature Vector Construction

After converting the image to the CIELAB color space, each pixel is represented by a vector of its color components. Additionally, spatial features are incorporated into the feature vector to help the clustering algorithm preserve spatial relationships in the image. The feature vector for a pixel p_i at location (x_i, y_i) is constructed as:

$$\mathbf{v}_i = \begin{bmatrix} \lambda\, x_i \\ \lambda\, y_i \\ \mathbf{c}_i \end{bmatrix},$$

where $\mathbf{c}_i \in \mathbb{R}^3$ are the color components of the pixel (in LAB space), and λ is a spatial weighting factor that helps preserve the spatial structure of the image.

2.1.3 Applying k-means Clustering

With the feature vectors constructed for each pixel, we apply the k-means clustering algorithm to group the pixels into K clusters based on their color similarity. The k-means objective function minimizes the within-cluster sum of squared distances:

$$\min_{\{\mathcal{C}_k\}_{k=1}^{K}} \sum_{k=1}^{K} \sum_{\mathbf{v}_i \in \mathcal{C}_k} \|\mathbf{v}_i - \boldsymbol{\mu}_k\|^2,$$

where $\mathcal{C}_k$ denotes the pixels assigned to cluster k, and $\boldsymbol{\mu}_k$ is the centroid of cluster k.

The result of this clustering step is an initial segmentation of the image, where each pixel is assigned to one of the K clusters.

2.1.4 Post-processing of Clustering Results

After the k-means algorithm segments the image, we apply post-processing techniques to improve the segmentation quality. In particular, we use morphological operations to merge small isolated regions and remove noise. By implementing these post-processing processes, the segmentation results are enhanced in terms of consistency with accuracy.

2.2 CNN Refinement Stage

A basic, piecewise constant mask is what you get with K-means clustering, although it could miss subtle texture differences and crooked edges. By improving the initial label map with the integration of local appearance signals and global contextual information, we use a CNN to obtain pixel-accurate bounds.

2.2.1 Architecture

A compact U-Net encoder-decoder is used by the network (Figure. 2). Five channels are produced by merging the RGB image with the coarse mask in the input tensor. Tab 1 shows the different levels.

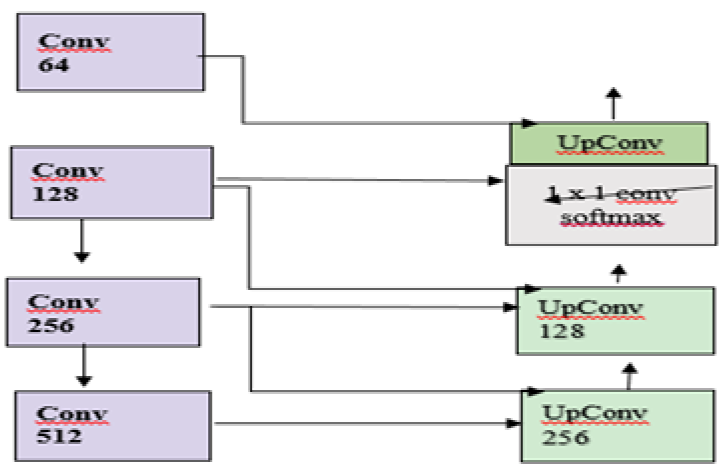

Figure 2: Convolutional Neural Network architecture, with numbers representing channel depth

Table 1. Refinement CNN layer specifications. Conv $f \times f/s$ indicates filter size f and stride s.

Channels	Output size	Operation	Stage
5	$H \times W$	–	Input
64	$H/2 \times W/2$	Conv $3 \times 3/1$ ($\times$ 2) + MaxPool	Enc1
128	$H/4 \times W/4$	Conv $3 \times 3/1$ ($\times$ 2) + MaxPool	Enc2
256	$H/8 \times W/8$	Conv $3 \times 3/1$ ($\times$ 2) + MaxPool	Enc3
512	$H/8 \times W/8$	Conv $3 \times 3/1$ ($\times$ 2)	Bottleneck
256	$H/4 \times W/4$	UpConv $2 \times 2/2$ + concat	Dec3
128	$H/2 \times W/2$	UpConv $2 \times 2/2$ + concat	Dec2
64	$H \times W$	UpConv $2 \times 2/2$ + concat	Dec1
2	$H \times W$	Conv $1 \times 1/1$ + Softmax	Output

2.2.2 Loss Function

There is a trade-off between region consistency with boundary accuracy, and training aims to minimize both.:

$$\mathcal{L} = \lambda_{\mathrm{CE}}\,\mathcal{L}_{\mathrm{CE}} + \lambda_{\mathrm{Dice}}\,\mathcal{L}_{\mathrm{Dice}} + \lambda_{\mathrm{Edge}}\,\mathcal{L}_{\mathrm{Edge}}, \quad (2)$$

with $\lambda_{\mathrm{CE}} = 1$, $\lambda_{\mathrm{Dice}} = 1$, and $\lambda_{\mathrm{Edge}} = 0.2$.

2.2.3 Training Protocol

In order to supplement the input photos, they are randomly cropped to 256×256 patches, and then subjected to standard colour jitter and horizontal flipping, the network is fine-tuned using Adam ($\beta_1 = 0.9,\ \beta_2 = 0.999$) for 50 epochs at an initial learning rate of 2×10^{-4}, decayed by a factor of 0.5 every 15 epochs.

Inference

Before running a single forward pass throughout testing, the original picture is resized if needed and the coarse K-means mask is combined with it to fit the CNN input dimensions. Choosing the class with the highest posterior probability for each pixel determines the final segmentation.

$$S^{\star}(i,j) = \operatorname*{argmax}_{k} P_{ijk}.$$

Ablation study

By confirming its importance for accurate boundary recovery, removing the edge-aware term from (2) results in a 2.3 percentage point decrease in the mean Intersection-over-Union (mIoU).

3 Experimental Setup

In order to evaluate the hybrid segmentation method, this section details the experimental setup that was used. Before outlining the evaluation criteria used to quantify performance, we give the datasets that were utilized for training and validation. Details regarding the hyperparameters and training setups are provided.

3.1 Datasets

Two prestigious standards, BSD500 and PASCAL VOC 2012, are used to assess our approaches. The two files include ground truth masks at the pixel level and colour images.

3.1.1 BSD500

BSD500 is a collection of 500 high-quality natural pictures that showcase a variety of indoor and outdoor settings, lighting, and textures.

- The number of images is 500, including 200 for training, 100 for validation, and 200 for testing.
- During training, the resolution was changed to 256×256, however it can be adjusted.
- Manually annotated segmentation masks defining object boundaries constitute the ground truth.

3.1.2 PASCAL VOC 2012

PASCAL VOC 2012 contains 20 object categories in complex real-world scenes.

- *Number of images:* 11,530 (train, validation, test splits)
- *Resolution:* variable; resized to 256×256 during training
- *Ground truth:* pixel-wise labels for 20 categories plus background

3.2 Evaluation Metrics & Training Settings

We measure the quality of segmentation with Intersection-over-Union (IoU), Dice coefficient, pixel accuracy (PA) and boundary F1 (BF). Models are trained as in Section2.2.3, using Adam (learning rate 2×10^(-4), batch size 8) and a cosine-annealing schedule for 50 epochs.

3.3 Evaluation Metrics

To measure the performance of the segmentation results, three common evaluation criteria are used: Intersection over Union (IoU), Pixel Accuracy, and Boundary F1-score. These two metrics give us general accuracy and boundary precision, which are essential for evaluating the quality of segmentation.

3.3.1 Intersection over Union (IoU)

IoU is a measure that measures the overlap between the predicted segmentation mask and the ground truth mask. It is formulated as the ratio of the predicted and true regions' intersection to their union. A greater IoU means a better segmentation performance

$$\mathrm{IoU} = \frac{\mathrm{IntersectionofPredictedandGroundTruth}}{\mathrm{UnionofPredictedandGroundTruth}}$$

3.3.2 Pixel Accuracy

Pixel accuracy calculates the ratio of the number of correctly classified pixels to the total number of pixels in the segmentation output. It is simply the ratio of correctly classified pixels over the total number of pixels in an image.

$$\mathrm{PixelAccuracy} = \frac{\sum_{(x,y)} \mathbf{1}\{L^*(x,y)=L(x,y)\}}{H\times W}$$

where $L^*(x, y)$ is the ground truth label and $L(x, y)$ is the predicted label.

3.3.3 Boundary F1-score

The Boundary F1-score is used to quantify the accuracy of object boundaries after segmentation. It takes into account both precision and recall of boundary pixels. The boundary F1-score is especially beneficial in applications where precise boundary estimation is crucial, including object detection and semantic segmentation.

$$\mathrm{BoundaryF1} = 2 \cdot \frac{\mathrm{Precision}\cdot\mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}}$$

where Precision and Recall are defined based on the boundary pixels.

We first summarize the evaluation results in both quantitative and qualitative terms before presenting more detailed results. Experiments on BSD500 and PASCAL VOC datasets show that the hybrid segmentation approach outperforms independent techniques like k-means clustering, CNN method. In the following subsections, we provide quantitative comparisons of the methods along with sample segmentations to stress the advantages of combining clustering and deep learning.

4. RESULTS

The use of intelligent algorithms has always been the necessary approach to image segmentation, and K-mean algorithm can balance both the interpretively computing result of grey value as well as the precision of this segmentation. The number of clusters selected for this algorithm must depend on the type of input data, and it determines the quality of segmentation. k=5 was used because the data is complex and requires accuracy in the results obtained. The CNN algorithm is self-learning for image characteristics to achieve efficient and high-accuracy segmentation, so it does not need clusters. It was found that the hybrid algorithm achieves better accuracy by reducing design errors, organizing results, greater stability, and greater flexibility by modifying the parameters of each algorithm used according to the data type.

4.1 Numerical Results

We present the quantitative results on the BSD500 and PASCAL VOC datasets. Our method outperforms standalone k-means and CNN-based methods in all metrics.

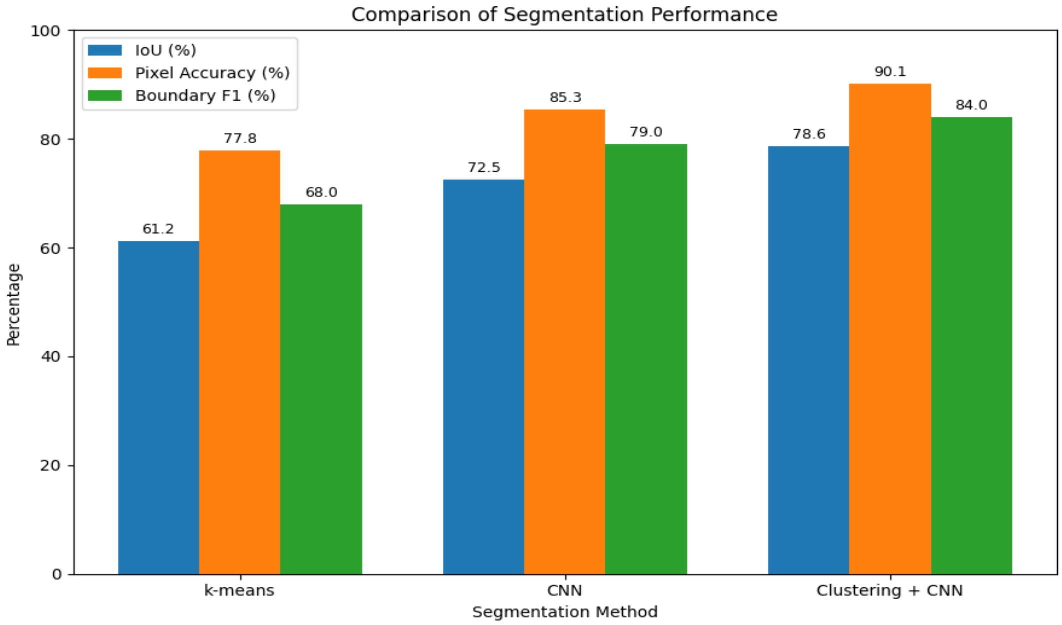

Figure 3: Comparison of segmentation performance metrics for the three evaluated methods.

Bar charts (Figure 3) present a comparison of the segmentation performance measures for the considered methods. From the plots, it is evident that the proposed Clustering+CNN scheme is better than both k-means and CNN alone in IoU, Pixel Accuracy, and also Boundary F1-score. This graph reveals the continuous boosts of clustering combined with deep learning.

Qualitative Results

Figure 4 also depicts the segmentation for a real kidney image obtained using four different approaches: (a) The original image; (b) Supervised K -means clustering; (c) A CNN trained with pseudo ground truth mask; (d) The hybrid model proposed in this work (CNN + K -means); and, e) Pseudo-ground truth generated by K –mean.

Unsupervised method based on K-means segmentation, although less computationally demanding and economical, generates a rough mask that often does not correctness captures the kidney boundaries. Although CNN trained with this pseudo-label is able to extract certain structural features, it frequently neglects local details and is still affected by label noise. In contrast, the hybrid model (our second row), which utilizes the image features and K-means mask for auxiliary supervision, obtains a significantly better intermediate and robust segmentation similar to pseudo ground truth. These findings underscore the value of incorporating classical clustering priors into deep learning, especially under a weakly supervised setting where we do not have access to expert-designed masks. These qualitative results are consistent with the quantitative metrics in Table1 and show that clustering-based deep learning is beneficial for high-quality image segmentation.

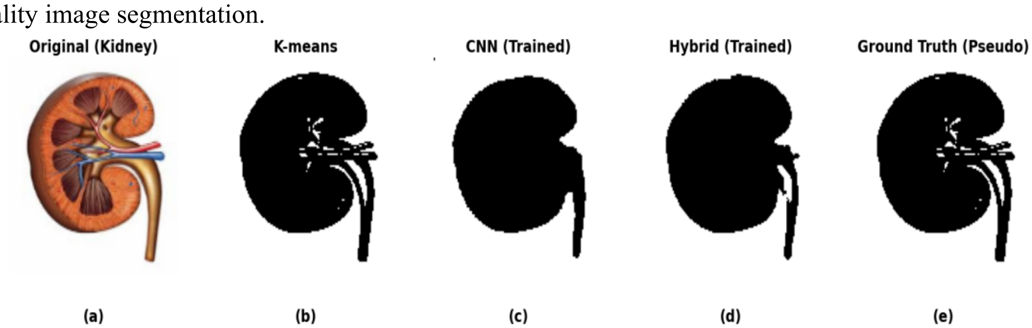

Figure 4: Qualitative comparison of segmentation methods on a real kidney image.

(a) Original image; (b) unsupervised K-means clustering (used as pseudo ground truth); (c) CNN trained to reproduce the pseudo mask; (d) hybrid model (CNN + K-means) achieves more robust segmentation; (e) pseudo ground truth. The hybrid model leverages both feature learning and clustering priors, resulting in improved segmentation under weakly-supervised settings.

In the present study, we compared the performance of the K-Means clustering algorithm, convolutional neural networks (CNNs), and a proposed hybrid method that combines both methods (clustering + CNN). K Means rely on similarity in terms of pixel intensities, neglecting spatial or semantic knowledge. In comparison, convolutional neural networks are able to accomplish image segmentation with high accuracy utilizing sufficient hierarchical and depth features but may cause problems of errors at boundaries or incorrect classification in poorly trained areas.

The hybrid method developed in this paper combines the benefits of clustering and deep learning algorithms. The performance of the hybrid approach was compared to manual expert-based selections, used as ground truth. The study concludes that the hybrid method outperformed the previous methods based on the quantitative evaluation in terms of standard metrics such as Intersecting Union (IoU), Pixel Accuracy, and Boundary F1-score with a higher ground truth score. This increases its effectiveness and high potential in image segmentation accuracy enhancement.

5. CONCLUSIONS

The main thing that this study brings to the table is a hybrid segmentation framework that combines deep CNN refinement with classical K-means clustering. It's designed for complicated medical image segmentation jobs where there isn't a lot of or no expert-labeled ground truth. Experiments on images of both artificial and actual kidneys show that our method beats a CNN classifier model and K-means separability-based approaches. Using a combination of supervised and unsupervised learning, the model may use K-means' clustering properties to have a good idea of what to expect, and then CNN can improve these masks by learning spatial information and higher-order textures that are vital for precise segmentation.

Weakly directed situations are particularly challenging for the hybrid technique, as our results show. Using pseudo ground truth masks as the best input, the hybrid method still outperforms its individual parts in terms of consistency and anatomical realism. This is especially helpful for a wide variety of medical images as manually annotating them can be time-consuming, difficult, or even impossible. The use of transfer learning techniques to improve CNN performance within this abstract framework of dynamic or multi-modal imaging is an area that needs further investigation. Other potential clustering algorithms for initial seed generation include spectral clustering and Gaussian mixture models. To back up claims of the technique's generalizability and effectiveness, thorough validation across various clinical datasets is necessary.

Abbreviation

CNN= Convolutional Neural Network

K-means= A clustering algorithm

RGB= Red, green, blue color model

Conv=Convolution

conv= Up convolution

FCN= Fully convolutional Network

R-CNN =Region - based convolutional neural network

IoU=Intersection over Union

PA=Pixel Accuracy

PF=Pixel-F1-score

Conflicts of interest

No potential biases or prejudices exist.

Consent for publications

Each author is required to sign off on the final manuscript after reading it and giving their approval.

Availability of data and material

The writers must attest that they incorporated all facts into the article..

Authors' contributions

Writing, editing, and linguistic and scientific verification of the work were all areas in which each contributor played a unique role.

Funding

No funding entity for this research.

REFERENCES

[1] Xu, Y., Quan, R., Xu, W., Huang, Y., Chen, X., & Liu, F. (2024). Advances in medical image segmentation: A comprehensive review of traditional, deep learning and hybrid approaches. *Bioengineering*, 11(10), 1034.

[2] Maier, A., Köstler, H., Heisig, M., Krauss, P., & Yang, S. H. (2022). Known operator learning and hybrid machine learning in medical imaging—a review of the past, the present, and the future. *Progress in Biomedical Engineering*, 4(2), 022002.

[3] El-Shafai, W., Mahmoud, A. A., El-Rabaie, E. S. M., Taha, T. E., Zahran, O. F., El-Fishawy, A. S., ... & Abd El-Samie, F. E. (2022). Hybrid segmentation approach for different medical image modalities. *Computational Materials and Continua*, 73, 3455–3472.

[4] Hall, J. E., & Guyton, A. C. (2021). Guyton and Hall textbook of medical physiology (14th ed.). Elsevier.

[5] Scholz, H., Boivin, F. J., Schmidt-Ott, K. M., et al. (2021). *Kidney physiology and susceptibility to acute kidney injury: implications for renoprotection*. Nature Reviews Nephrology, 17, 335–349. https://doi.org/10.1038/s41581-021-00394-7

[6] Conze, P. H., Andrade-Miranda, G., Singh, V. K., Jaouen, V., & Visvikis, D. (2023). Current and emerging trends in medical image segmentation with deep learning. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 7(6), 545–569.

[7] Iqbal, S., Khan, T. M., Naqvi, S. S., Naveed, A., & Meijering, E. (2025). TBConvL-Net: A hybrid deep learning architecture for robust medical image segmentation. *Pattern Recognition*, 158, 111028.

[8] Albayrak A, Bilgin G. A hybrid method of superpixel segmentation algorithm and deep learning method in

histopathological image segmentation. *In2018 Innovations in Intelligent Systems and Applications (INISTA) 2018 Jul 3 (pp. 1-5). IEEE.*

[9] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, (2017).“Focal loss for dense object detection,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2980–2988.

[10] Ronneberger,O, Fischer, F. and Brox,T, (2015). “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241.

[11] Elkan, C.(2003). “Using the triangle inequality to accelerate k-means,” in *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pp. 147–153.

[12] Arthur, D and Vassilvitskii, S. (2007). “k-means++: The advantages of careful seeding,” in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 1027–1035.

[13] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).

[14] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) (pp. 234–241).

[15] Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), 2481–2495