Pancreas CT Segmentation by Predictive Phenotyping

Yucheng Tang[1(✉)], Riqiang Gao[1], Hohin Lee[1], Qi Yang[1], Xin Yu[1], Yuyin Zhou[2], Shunxing Bao[1], Yuankai Huo[1], Jeffrey Spraggins[1,5], Jack Virostko[3], Zhoubing Xu[4], and Bennett A. Landman[1,5]

[1] Vanderbilt University, Nashville, TN 37203, USA
yucheng.tang@vanderbilt.edu
[2] Stanford University, Stanford, CA 94305, USA
[3] University of Texas at Austin, Austin, TX 78705, USA
[4] Siemens Healthineers, Princeton, NJ 08540, USA
[5] Vanderbilt University Medical Center, Nashville, TN 37235, USA

Abstract. Pancreas CT segmentation offers promise at understanding the structural manifestation of metabolic conditions. To date, the medical primary record of conditions that impact the pancreas is in the electronic health record (EHR) in terms of diagnostic phenotype data (*e.g.*, ICD-10 codes). We posit that similar structural phenotypes could be revealed by studying subjects with similar medical outcomes. Segmentation is mainly driven by imaging data, but this direct approach may not consider differing canonical appearances with different underlying conditions (*e.g.,* pancreatic atrophy versus pancreatic cysts). To this end, we exploit clinical features from EHR data to complement image features for enhancing the pancreas segmentation, especially in high-risk outcomes. Specifically, we propose, to the best of our knowledge, the first phenotype embedding model for pancreas segmentation by predicting representatives that share similar comorbidities. Such an embedding strategy can adaptively refine the segmentation outcome based on the discriminative contexts distilled from clinical features. Experiments with 2000 patients' EHR data and 300 CT images with the healthy pancreas, type II diabetes, and pancreatitis subjects show that segmentation by predictive phenotyping significantly improves performance over state-of-the-arts (Dice score 0.775 to 0.791, $p < 0.05$, Wilcoxon signed-rank test). The proposed method additionally achieves superior performance on two public testing datasets, BTCV MICCAI Challenge 2015 and TCIA pancreas CT. Our approach provides a promising direction of advancing segmentation with phenotype features while without requiring EHR data as input during testing.

Keywords: Pancreas segmentation · Patient phenotype · Predictive clustering

1 Introduction

Patient care data, such as CT scans and electronic health records (EHR) with the pancreatic disease, are heterogeneous in nature. Disease progression and treatment delivery are associated with different care trajectories, which in turn lead to varying pancreas patterns. In anticipation of diabetes, patients are observed with atrophic pancreas

M. de Bruijne et al. (Eds.): MICCAI 2021, LNCS 12901, pp. 25–35, 2021.
https://doi.org/10.1007/978-3-030-87193-2_3

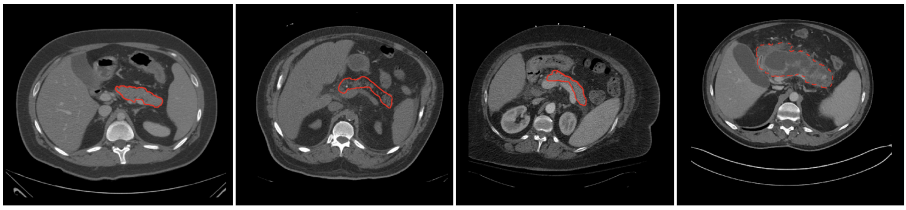

Fig. 1. Representative images are predicted to associate with comorbidities and ICD-10 codes (phenotype components) identified in each risk category. The red outlines show the pancreas tissue can be different under phenotyping contexts. (1) is from a nominally healthy pancreas group with potential lung infections; (2) is from type I diabetes and other chronic kidney disease patients with atrophic pancreas; (3) is from other metabolic syndromes including type II diabetes; (4) is from patients with weight loss and pancreatitis.

tissues [5], with progression noted in the patients' medical history in terms of International Classification of Diseases (ICD) codes. Current pancreas segmentation methods [15,18,24,25] are typically driven by imaging data, while phenotype covariates [3,6,11,13,22] that indicate underlying patient conditions are not well considered. We observe that different disease types present heterogeneous textures (Fig. 1), and thereby hypothesize that identifying different pancreas patterns can extract the discriminative contexts which can well benefit pancreas segmentation.

Data-driven phenotype clustering has been recently used to group patients sharing close outcomes [1,4,8,9]. Combining imaging biomarkers, Virostko *et al.* [20] assessed the pancreas size with type I diabetes patients. Tang *et al.* [19] showed the feasibility of onset type II diabetes prediction using CT scans. However, to date, how to fully exploit the EHR data for guiding medical image segmentation has been rarely studied. A naïve approach is to simply concatenate both image and EHR data as a two-channel input, and then train a standard convolutional neural network for deriving the outcome. However, this fusion strategy is not directly applicable for our task since 1) a patient can have hundreds of phenotype categories; 2) the fusion strategy cannot account for patients' observed outcomes (*e.g.*, onset of comorbidities, chronic progression of metabolic syndrome); and 3) it requires EHR data as input during inference, which does not commonly exist for many real-world pancreas segmentation datasets (*e.g.*, BTCV MICCAI Challenge [7], TCIA pancreas CT [14]).

To address above challenges, we propose the first pancreas segmentation framework to model both the pancreas imaging features and clinical features via predictive phenotyping. The rationale is that the larger scale of EHR data with (*e.g.*, ICD-10 code) which indicates phenotype subgroups can be potentially correlated to the different appearance of the pancreas. Specifically, the proposed approach consists of an encoder, a segmentation decoder, and a predictor with sets of phenotypes candidates' centroids. Our method

is designed to meet the following requirements: 1) The subject image should be partitioned into several subgroups sharing similar future outcomes; 2) The assigned discrete representation should retain the patient phenotype context, and 3) The phenotype representation is used as prior knowledge for predicting. Particularly, in our framework, the encoder maps an image into a latent representation; the predictor assigns one or several phenotype categories by taking the latent variable as input; the segmentation decoder estimates the pixel-wise labels conditioned on the assigned centroid. To homogenize future outcomes in each subgroup, we introduce a phenotyping objective given CT images by regularizing Kullback-Leibler (KL) divergence between the learned latent representation and the embedding centroid. Finally, the segmentation model estimates the pancreas segmentation mask given encoding of the image and the risk embedding.

Our contribution is four-folds: we successfully (1) learn a phenotype embedding between CT and EHR; (2) formulate a pancreas segmentation framework that benefits from predicting phenotype subgroups; (3) demonstrate improved pancreas segmentation performance on healthy and disease patient cohorts; (4) design the embedding approach without requiring EHR at the testing phase. The significance of the study is that we use an experimental CT imaging-phenotyping approach for investing clinical underpinnings of pancreas segmentation. The phenotype embedded model enriches segmentation contexts improving the characterization of heterogeneous disease and allows for deeper consideration of patient phenotype in image-based learning.

2 Method

2.1 Problem Formulation

Let $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ be variables for input images and an output segmentation label. $C \in \mathcal{C}$ is the patient phenotype onset (*i.e.*, one or a combination of future outcomes) where $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{C}$ are the image feature, label, and phenotype onset space, respectively. Specifically, CT image X selection is censored from timestamp: date of diagnostic code is later than date of scan at least 1 year. The input of C is the sequence of covariate admissions, the feature of one admission is a multi-hot vector containing the comorbidities or demographics: $C'_{EHR} = [M', M_1, M_2, ..., M_{\mathcal{C}}]$. M' is the set of demographic values, M_C is the set of phenotype admissions constructed by binary vectors of aggregate ICD-10 codes. In the training phase, we are given the dataset $\mathcal{D} = \{x^n, y^n, c^n\}_{n=1}^N$ consists of observations (x, y, c) for N subjects. In the testing phase, we assume the dataset only comprising image volume $\{x^n\}_{n=1}^N$. The goal is first to identify a set of K predictive phenotypes $\mathcal{Z} = \{z_1, z_2, ..., z_K\}$ lying in the latent space. Each phenotype cluster is supposed comprising of homogeneous patients that can be represented by the cluster centroid. This predictive phenotyping updates the encoder to suggest the context to which cluster a patient belongs. Second, we design the segmentation model to estimate the pixel-wise label given the encoding variable and the predictive phenotyping distribution. Let $\bar{z}$ be the random variable that lying in the phenotype onset latent space and s be the image feature. The predictive phenotyping can be fulfilled by optimizing the Kullback-Leibler (KL) divergence between distributions conditioned on the image:

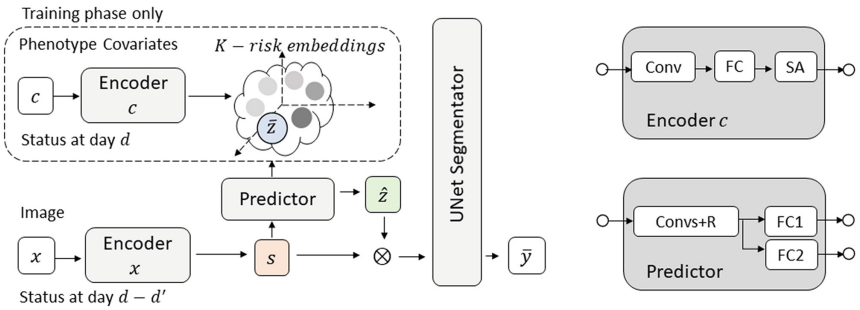

Fig. 2. Phenotype embedded segmentation architecture. The left diagram shows the embedding network combining image features s, predictive phenotyping $\hat{z}$, pre-existing risk conditions $\bar{z}$ lying in the latent space to be fed into the segmentation model. The predictor is trained for predicting phenotype-dependent feature maps and selecting "similar" cluster assignment, where the phenotype information is not required as input in the testing phase. Right top: encoder for processing phenotype covariates. Right bottom: the predictor follows the self-training scheme for image feature. Here, SA denotes soft assignment for risk embeddings, R for ReLU, FC for fully connected layers, and $\otimes$ denotes concatenation.

$p(\hat{z}|s)$ and onset phenotypes $p(\bar{z}|c)$, respectively. Combining the aim of segmentation, we establish our goal as following objective:

$$\underset{Y}{\text{minimize}}\, \mathbb{E}_{Y\sim(x,y,c)} \left[-\log P(y|s,\hat{z}_k)\right] + KL\left(\hat{z}_k|s \parallel \bar{z}_k|c\right). \tag{1}$$

2.2 Loss Functions

Loss functions are designed to meet the objective in Eq. 1 and are proposed to iteratively refine the predicted phenotyping from image features. Specifically, our model is trained by matching image distribution to the target distribution defined by future outcomes. To this end, we define the objective as KL divergence between an expectation and the cluster assignment:

$$\mathcal{L}_1(\bar{z},\hat{z}) = \mathbb{E}_{z\sim P(x,c)}\left[-\sum_{k=1}^{\mathrm{K}} \bar{z}_k \log \hat{z}_k\right], \tag{2}$$

where $\bar{z}_k$ and $\hat{z}_k$ indicate the k-component of $\bar{z}$, $\hat{z}$, respectively. Note that the KL divergence loss reaches its minimum when two latent distributions are equivalent. Additionally, the segmentation loss penalizes the predicted mask $\bar{y}$ and the ground truth label y by DSC-loss:

$$\mathcal{L}_2(\bar{y},y) = \mathbb{E}_{Y\sim P(s,\hat{z})}\left[1 - \frac{2\times\sum_i y_i\bar{y}_i}{\sum_i y_i + \sum_i \bar{y}_i}\right], \tag{3}$$

where the form follows [23] to prevent a model from background bias.

2.3 Phenotype Embedding

To encourage homogeneous future outcomes in each phenotyping cluster, we employ embedded mapping [21] as our initialization method. Given an initial estimate of the non-linear mapping c' and cluster centroid μ. We adopt the self-supervision [12] training strategy that iteratively 1) optimizes soft-assignment between embedded points and clustering centroids; 2) updates deep mapping and centroids. The soft-assignment block in Fig. 2 follows [10] using Student's t-distribution as a kernel to estimate between data points and cluster centroid:

$$q_{ik} = \frac{\exp\left(1+\left\|c'_i-\mu_k\right\|^2/\alpha\right)}{\sum_j \exp\left(1+\left\|c'_i-\mu_j\right\|^2/\alpha\right)}\ , \tag{4}$$

where α denotes the degrees of freedom of Student's t-distribution ($\alpha = 1$ for all experiments), $\exp$ is the exponential operation with power $-(\alpha+1)/2$ and q_{ik} can be interpreted as the probability of assigning sample i to cluster k. More comparisons of clustering benchmarks can be found in [8,9]. After initialization, the embedding learning is iteratively updated during segmentation training.

3 Experiments

3.1 Dataset

The Abnormal Pancreas Segmentation Dataset. We have curated an abnormal pancreas dataset that contains 2000 adult patients (aged 18–50 years) with 14927 recorded visits and de-identified longitudinal CT scans under IRB approval. Each patient is associated with 101 covariates under radiologists' query, including information on demographic and abdomen-related comorbidities that can potentially impact pancreas tissues. CT images are acquired at least one year (range from 1.0 to 2.1 years) earlier than diagnosis codes for each patient, to meet the requirements of the prognostic task with predictive phenotyping. For the segmentation task, 300 patients' CT images are annotated and used for experiments. Each CT scan is 512 × 512 × Slices, where the number of slices ranges from 72 to 121 under the body part regression process of [17] to acquire relatively same abdomen region of interest (ROI). The slice thickness ranges from 1mm to 2.5mm.

BTCV MICCAI Challenge 2015. We used the MICCAI 2015 Multi-Atlas Abdomen Labeling Challenge [7] as one of the external testing sets. The challenge dataset contains 50 abdominal CT scans. For evaluating the testing phase of the proposed method, the dataset does not include patient phenotype information. Each CT scan is manually labeled with 13 structures including pancreas with a spatial resolution of ([0.54 ∼ 0.54] × [0.98 ∼ 0.98] × [2.5 ∼ 5.0] mm^3).

TCIA Pancreas CT. We use the 82 abdominal contrast enhanced CT scans from National Institutes of Health Clinical Center as the second external testing set. The publicly available study cohort contains 17 kidney donor subjects, and 65 patients were selected with no pancreatic cancer lesions and pathology. Each CT scan is in a resolution of 512 × 512 and slice thickness of [1.5 ∼ 2.5] mm.

3.2 Implementation Details

Follow prevailing pancreas segmentation baselines [18,24,25], we adopt the coarse-to-fine strategy for 3D pancreas segmentation. The coarse stage takes a highly down-sampled CT volume at an input dimension of $164 \times 164 \times 64$. For the fine stage, we cropped $64 \times 64 \times 64$ sub-volumes constrained to be in the pancreas region of interest (ROI). For experiments, 10% and 20% of subjects are randomly selected as validation and testing sets with the in-house dataset. Note that, the two external datasets are only used for testing, no subjects are used for the training procedure. We used 1) CT window range of [-175, 275] HU; 2) scaled intensities of [0.0,1.0]; 3) training with Nvidia 2080 11GB GPU with Pytorch implementation; 4) Adam optimizer with momentum 0.9. The Learning rate is initialized to 0.001 followed by a factor of 10 every 50 epochs decay.

Metrics. Segmentation performance is evaluated between ground truth and prediction by Dice-Sorensen coefficient (DSC), Averaged Surface Distance (ASD), and symmetric Hausdorff Distance (HD).

3.3 Comparison with State-of-the-Arts

We compare the proposed method with various state-of-the-art methods: 1) 3D-UNet [2]; 2) hierarchical 3D FCN [16] (denoted as "3D FCN"); 3) the fixed-point model [24] (denoted as "C2F Fixed-point"); 4) 3D ResDSN [25] (denoted as "C2F ResDSN"); and 5) the random patches model [18] (denoted as "C2F Random-patches"). Here "C2F" denotes the coarse-to-fine training strategies [18,24,25].

3.4 Results

We compare our method against state-of-the-art approaches with respect to the cluster number at 4 (Table 1). Our method significantly improves performance in terms of DSC, ASD and HD, with $p < 0.05$ under Wilcoxon signed-rank test. Importantly, the Hausdorff distance (HD) improvement shows that EHR information provided useful context to reduce outliers. In Table 2, we further investigate the comparison experiment results with external testing sets. The two public challenge data do not include patient EHR, *i.e.*, demographics, ICD codes. Our method implicitly predicts the future outcomes from the image feature and fused to the segmentation task. The method achieves a mean DSC of 0.757 on BTCV data, and 0.827 on TCIA pancreas CT. Predictive phenotyping improves several outlier cases, showing less variance (Fig. 3). Qualitative inspection confirms the numerical results (Fig. 4). First, we inspect the data of a patient with potential lung infections and relatively normal pancreas tissue. In the second case, the patient has type I diabetes, observing a degraded pancreas tissue. Importantly, the improvement with respect to the degraded pancreas is larger than the healthy pancreas, showing the predictive phenotyping can be informative for identifying variant patterns.

Ablative Study. *Efficacy of the predictive phenotyping and network architecture* In Table 1, we compared the backbone model (row 5) and predictive phenotyping (row 6). The EHR improved performance on two datasets by 1.5%, significant Wilcoxon signed-rank test, $p<0.001$. For the diabetic patients, performance improvement gains

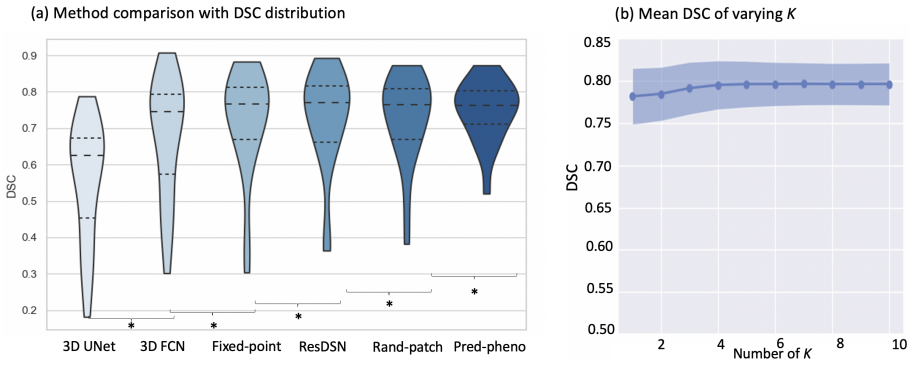

Fig. 3. Testing performance on the in-house dataset. Left: Distribution (median and quartiles) of DSC, the predictive phenotyping shows smaller variance and reduces the number of outliers (DSC < 0.4). Right: The DSC (mean) comparison with varying K. The performance shows higher improvement as K increases from 1 to 4, then becomes marginal after $K = 4$. * denotes statistically significant under Wilcoxon signed-rank test ($p < 0.05$).

Table 1. Performance comparison on the abnormal pancreas segmentation dataset. C2F denotes coarse-to-fine training. * denotes statistically significant against above method with Wilcoxon signed-rank test.

Methods	DSC	ASD	HD
3D-UNet (Cicek *et al.*)	0.697	5.592	27.154
3D FCN (Holger *et al.*)	0.724*	4.042*	25.195*
C2F Fixed-point (Zhou *et al.*)	0.746*	2.981*	22.516*
C2F ResDSN (Zhu *et al.*)	0.767*	2.105	22.017
C2F Random-patches (Tang *et al.*)	0.775*	1.976*	20.591*
Predictive phenotyping (Ours, K = 4)	**0.791***	**1.697***	**19.482***

were larger. Predictive phenotyping with the EHR outperforms naïve approach with feature concatenation by a large margin, from 74.5% to 77.9% (Fig. 3). In Table 1 and Fig. 3, we compared with pancreas segmentation state-of-the-art methods. Predictive phenotyping significantly improved performance in terms of DSC, ASD and HD, with $p<0.05$, Wilcoxon signed-rank test. Importantly, HD improvement shows that EHR information provided useful context to reduce outliers. In Table 2, we further investigate the comparison experiment results with external testing sets. For external validation, the two public challenge data do not include patient EHR, i.e., demographics, ICD codes. Our method implicitly predicts the future outcomes from the image feature and fused to the segmentation task. The method achieves a mean DSC of 0.757 on BTCV data, and 0.827 on TCIA pancreas CT.

Importance of hyper-parameter K. We further evaluate the performance by varying the number of clusters K from 1 to 10 on the in-house dataset. Figure 3 shows improved

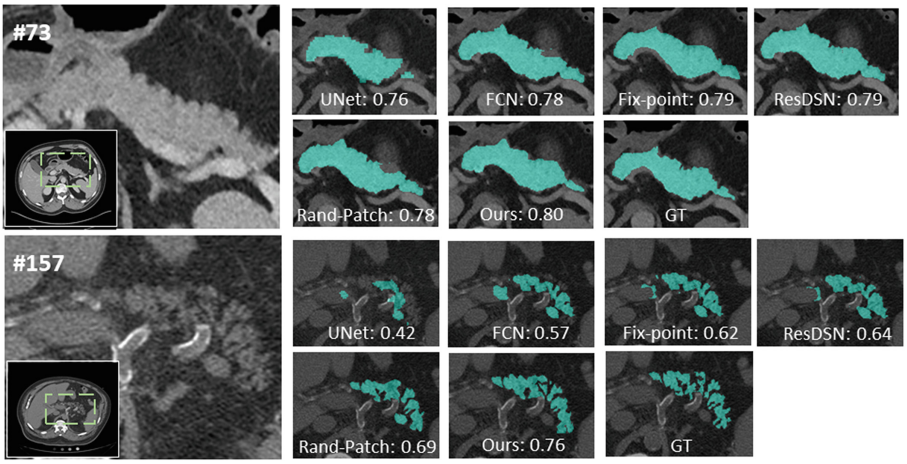

Fig. 4. Two representative cases. The top subject has potential lung infections and relative normal pancreas tissue. The bottom case has type I diabetes with degraded pancreas tissue. The accuracy gain of the diabetes case is larger than the normal case, showing the method's ability for identifying variant morphological pancreas.

Table 2. External testing performance comparison on BTCV MICCAI Challenge 2015 and TCIA pancreas (mean DSC) with our model trained on the internal data. Note that no subject from these two datasets are used for training. C2F denotes coarse-to-fine training strategies. * for statistically significant against above method with Wilcoxon signed-rank test.

Methods	BTCV	TCIA
3D-UNet (Cicek *et al.*)	0.685	0.770
3D FCN (Holger *et al.*)	0.709*	0.776*
C2F Fixed-point (Zhou *et al.*)	0.726*	0.797*
C2F ResDSN (Zhu *et al.*)	0.730*	0.804
C2F Random-patches (Tang *et al.*)	0.742*	0.813*
Predictive phenotyping (Ours, K = 4)	**0.757***	**0.822***

DSC as K increased, the DSC improves from 0.7814 to 0.7956 as K from 1 to 4. The performance is observed no significant improvement after $k = 4$ ($p < 0.1$, Wilcoxon signed-rank test).

4 Discussion and Conclusion

Comparing Table 1 and Table 2, the proposed method shows higher improvement over baseline methods if the cohort has more severe cases of abdominal diseases. Specifically, the performance improvement on the abnormal pancreas segmentation dataset in terms of the average Dice is 1.6%, which is larger than that of the BTCV dataset (1.5%), and the TCIA dataset (1.1%), respectively. We have also demonstrated two qualitative

examples in Fig. 4, to show that our method can lead to more performance gain for the atrophic pancreas than the normal pancreas. The larger improvement on diseased cohort can be a potential advantage of the phenotype embedding. In addition to the major segmentation objective, the case-specific feature projected to the phenotype embedding space can be observed in Fig. 1. The comorbidities developed in the next two years with 4 identified clusters, and listed ICD-10 codes are with most frequencies in each grouped phenotype component. The first component shares the most cases with relative normal pancreas, while the second, third and fourth indicate varying phenotype outcomes of the atrophic pancreas, metabolic syndrome, and pancreas with inflammatory fats, respectively. The number of phenotype components K is one of the most important parameters in the study: increasing k can potentially impact the predictive embedding with higher diversity representing data distribution. However, the interpretability will decrease as it shares fewer similar data points. In the future, the interpretability of the predicted phenotyping can be further evaluated with more clinically meaningful investigations.

In this work, we introduce pancreas segmentation by predictive phenotyping, a patient-oriented approach for understanding between EHR and CT data. The experimental imaging-phenotyping approach is used for investigating the phenotype underpinnings of the pancreas. We demonstrate a predictive task to encourage image embedding to the phenotyping cluster with similar patient outcomes. The EHR data is designed as input at the training phase, and only images are required for inferencing phenotyping context and segmentation at the test phase. Throughout experiments on the in-house dataset and two public challenge datasets, we show that the method highlights a significant role over state-of-the-art segmentations. The integrated imaging-phenotyping method could encourage solutions that better respect anatomical variability, especially associated with disease progression or comorbidities. When EHR data is available, the method can be applied for boosting performance.

Acknowledgements. This research is supported by NIH Common Fund and National Institute of Diabetes, Digestive and Kidney Diseases U54DK120058, NSF CAREER 1452485, NIH grants, 2R01EB006136, 1R01EB017230 (Landman), and R01NS09529. The identified datasets used for the analysis described were obtained from the Research Derivative (RD), database of clinical and related data. The imaging dataset(s) used for the analysis described were obtained from ImageVU, a research repository of medical imaging data and image-related metadata. ImageVU and RD are supported by the VICTR CTSA award (ULTR000445 from NCATS/NIH) and Vanderbilt University Medical Center institutional funding. ImageVU pilot work was also funded by PCORI (contract CDRN-1306-04869).

References

1. Baytas, I.M., Xiao, C., Zhang, X., Wang, F., Jain, A.K., Zhou, J.: Patient subtyping via time-aware lstm networks. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 65–74 (2017)
2. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 424–432. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49

3. Evans, J.A.: Electronic medical records system (Jul 13 1999), uS Patent 5,924,074
4. Giannoula, A., Gutierrez-Sacristán, A., Bravo, Á., Sanz, F., Furlong, L.I.: Identifying temporal patterns in patient disease trajectories using dynamic time warping: a population-based study. Sci. Rep. **8**(1), 1–14 (2018)
5. Goda, K., Sasaki, E., Nagata, K., Fukai, M., Ohsawa, N., Hahafusa, T.: Pancreatic volume in type 1 und type 2 diabetes mellitus. Acta Diabetol. **38**(3), 145–149 (2001)
6. Hales, C.N., Barker, D.J.: Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. Diabetologia **35**(7), 595–601 (1992)
7. Landman, B., Xu, Z., Igelsias, J., Styner, M., Langerak, T., Klein, A.: Miccai multi-atlas labeling beyond the cranial vault-workshop and challenge. In: Proceedings of MICCAI Multi-Atlas Labeling Beyond Cranial Vault—Workshop Challenge (2015)
8. Lee, C., Van Der Schaar, M.: Temporal phenotyping using deep predictive clustering of disease progression. In: International Conference on Machine Learning, pp. 5767–5777. PMLR (2020)
9. Luong, D.T.A., Chandola, V.: A k-means approach to clustering disease progressions. In: 2017 IEEE International Conference on Healthcare Informatics (ICHI), pp. 268–274. IEEE (2017)
10. Van der Maaten, L., Hinton, G.: Visualizing data using t-sne. J. Mach. Learn. Res. **9**(11) (2008)
11. Mani, S., Chen, Y., Elasy, T., Clayton, W., Denny, J.: Type 2 diabetes risk forecasting from emr data using machine learning. In: AMIA Annual Symposium Proceedings, vol. 2012, p. 606. American Medical Informatics Association (2012)
12. Misra, I., Maaten, L.v.d.: Self-supervised learning of pretext-invariant representations. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6707–6717 (2020)
13. Quan, H., et al.: Coding algorithms for defining comorbidities in icd-9-cm and icd-10 administrative data. Medical care, pp. 1130–1139 (2005)
14. Roth, H., Farag, A., Turkbey, E., Lu, L., Liu, J., Summers, R.: Data from pancreas-ct. The cancer imaging archive (2016)
15. Roth, H.R., Farag, A., Lu, L., Turkbey, E.B., Summers, R.M.: Deep convolutional networks for pancreas segmentation in CT imaging. In: Medical Imaging 2015: Image Processing, vol. 9413, p. 94131G. International Society for Optics and Photonics (2015)
16. Roth, H.R., et al.: DeepOrgan: multi-level deep convolutional networks for automated pancreas segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9349, pp. 556–564. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24553-9_68
17. Tang, Y., et al.: Body part regression with self-supervision. IEEE Trans. Med. Imaging **40**, 1499–1507 (2021)
18. Tang, Y., et al.: High-resolution 3d abdominal segmentation with random patch network fusion. Med. Image Anal. **69**, 101894 (2021)
19. Tang, Y., et al.: Prediction of Type II diabetes onset with computed tomography and electronic medical records. In: Syeda-Mahmood, T., et al. (eds.) CLIP/ML-CDS -2020. LNCS, vol. 12445, pp. 13–23. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-60946-7_2
20. Virostko, J., Hilmes, M., Eitel, K., Moore, D.J., Powers, A.C.: Use of the electronic medical record to assess pancreas size in type 1 diabetes. PLoS ONE **11**(7), e0158825 (2016)
21. Xie, J., Girshick, R., Farhadi, A.: Unsupervised deep embedding for clustering analysis. In: International Conference on Machine Learning, pp. 478–487. PMLR (2016)
22. Zheng, T., et al.: A machine learning-based framework to identify type 2 diabetes through electronic health records. Int. J. Med. Informatics **97**, 120–127 (2017)

23. Zhou, Y., Xie, L., Fishman, E.K., Yuille, A.L.: Deep supervision for pancreatic cyst segmentation in abdominal CT scans. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 222–230. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66179-7_26
24. Zhou, Y., Xie, L., Shen, W., Wang, Y., Fishman, E.K., Yuille, A.L.: A fixed-point model for pancreas segmentation in abdominal CT scans. In: MICCAI, pp. 693–701 (2017)
25. Zhu, Z., Xia, Y., Shen, W., Fishman, E., Yuille, A.: A 3d coarse-to-fine framework for volumetric medical image segmentation. In: 2018 International Conference on 3D Vision (3DV), pp. 682–690. IEEE (2018)