Engineering 7 (2021) 881–888

Contents lists available at ScienceDirect

Engineering

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

Research
Intelligent Medicine—Article

# Multi-View Point-Based Registration for Native Knee Kinematics Measurement with Feature Transfer Learning

Cong Wang [a,b,c,#], Shuaining Xie [b,#], Kang Li [d], Chongyang Wang [e], Xudong Liu [e,*], Liang Zhao [b,*], Tsung-Yuan Tsai [a,b,c,*]

[a] Shanghai Key Laboratory of Orthopaedic Implants & Clinical Translational R&D Center of 3D Printing Technology, Department of Orthopaedic Surgery, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine; School of Biomedical Engineering & Med-X Research Institute, Shanghai Jiao Tong University, Shanghai 200030, China
[b] SenseTime Research, Shanghai 200233, China
[c] Engineering Research Center of Digital Medicine and Clinical Translation, Ministry of Education, Shanghai 200030, China
[d] Department of Orthopaedics, New Jersey Medical School, Rutgers University, Newark, NJ 07103, USA
[e] Department of Orthopaedics, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai 200233, China

ARTICLE INFO

*Article history:*
Received 17 February 2020
Revised 22 March 2020
Accepted 30 March 2020
Available online 22 July 2020

*Keywords:*
2D–3D registration
Machine learning
Domain adaption
Point correspondence

ABSTRACT

Deep-learning methods provide a promising approach for measuring *in-vivo* knee joint motion from fast registration of two-dimensional (2D) to three-dimensional (3D) data with a broad range of capture. However, if there are insufficient data for training, the data-driven approach will fail. We propose a feature-based transfer-learning method to extract features from fluoroscopic images. With three subjects and fewer than 100 pairs of real fluoroscopic images, we achieved a mean registration success rate of up to 40%. The proposed method provides a promising solution, using a learning-based registration method when only a limited number of real fluoroscopic images is available.



## 1. Introduction

Accurate kinematics of the knee joint is critical in many orthopedic applications for understanding aspects such as the normal function of the joint [1], development of knee osteoarthritis [2], mechanisms of knee injuries [3], optimization of prosthesis design [4], preoperative planning, and postoperative rehabilitation [5]. The measurement of knee kinematics is also essential for biomechanical studies on the musculoskeletal system. In the event of significant demand for kinematics in the clinical field, an efficient and reliable method to measure the dynamic motion of the joint is needed.

Various measurement tools are now available for researchers to quantify three-dimensional (3D) knee kinematics, but only a few of them can provide millimeter-scale accuracy and rapid tracking velocity. Skin-marker-based optical tracking systems are widely used in the analysis of human motion, but their accuracy is affected by marker-associated soft-tissue artifacts, which can cause displacements of up to 40 mm [6]. Although several researchers have attempted to reduce the effects of soft-tissue artifacts by building mathematical models [7–9], the issue remains unsolved when using any skin-marker-based motion-capture technique [10]. With the development of medical imaging, some techniques can measure dynamic joint kinematics directly, such as real-time magnetic resonance (MR) tomography and computed tomography (CT) [11,12]. However, clinical promotion of these techniques was limited by low temporal resolution, restricted range of motion (ROM), the need to control motion speed, low image quality, and non-negligible amounts of radiation [13,14]. In the past decade, a dual-fluoroscopic imaging system (DFIS) has been widely used and well-accepted for accurate *in-vivo* joint motion analysis because of its high accuracy [15], accessibility, sufficient ROM [16], and low radiation levels compared with traditional radiography (Fig. 1).

To find the pose of the object (i.e., native knee joints) in DFIS, two-dimensional (2D) to 3D registration, which aligns the volume data (e.g., CT) with fluoroscopy (continuous X-ray images), is applied in the measurement procedure. The 3D position of CT is adjusted iteratively, and a large number of digitally reconstructed

* Corresponding authors.
*E-mail addresses:* xdliu@sjtu.edu.cn (X. Liu), zhaoliang@sensetime.com (L. Zhao), tytsai@sjtu.edu.cn (T.-Y. Tsai).
[#] These authors contributed equally to this work.

https://doi.org/10.1016/j.eng.2020.03.016

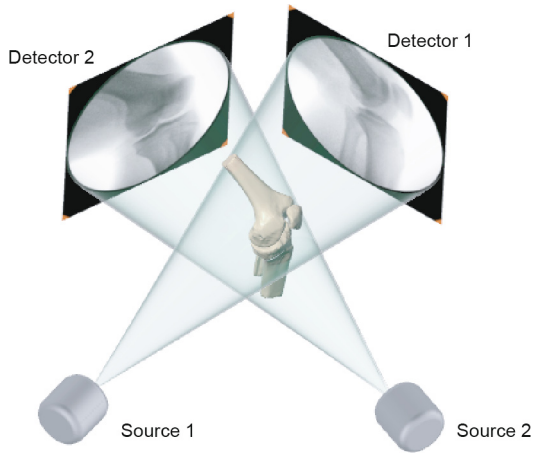


**Fig. 1.** Virtual DFIS for measuring the dynamic motion of knee joints.

radiographs (DRRs) is generated simultaneously until the DRR is most similar to the X-ray image [17]. With the increasing use of DFIS in clinical applications, researchers have attempted various automatic registration methods to accelerate the matching procedure. Optimization-based registration, which is composed of an optimizer and similarity metrics between images, has been investigated extensively [18,19]. Although the accuracy of optimization-based registration is high [20–22], several drawbacks, such as the strictly required registration initialization and the high computational cost of calculating DRRs and the iterations during optimization, limit the widespread use of DFIS [23].

With the rapid development of machine learning [24,25] in recent years, several learning-based methods have been developed to measure joint kinematics, with the advantages of computational efficiency and enhancement of capture range compared with optimization-based methods [21,26–28]. However, these methods are always trained with synthetic X-ray images (i.e., DRRs) because training such models with a large amount of authentic labeled data is impractical. Even so, considerable authentic images are still necessary to ensure the robustness of registration [22,27]. Another consideration is the discrepancy between DRRs and X-ray images. Compared with DRRs, fluoroscopic images showed blurred edges, geometric distortion, and nonuniform intensity [29,30]; therefore, networks that train on DRRs do not generalize to fluoroscopic images ideally [22]. Previous studies have established various physical models to generate more realistic DRRs through additional measurements of X-ray quality [31,32]. Recently, a survey conducted by Haskins et al. [24] has shown the ability to use transfer learning in such cross-modal registration, which may save the effort of building complicated DRR models or collecting authentic clinical images.

In our work, we developed a pseudo-Siamese multi-view point-based registration framework to address the problem of limited number of real fluoroscopic images. The proposed method is a combination of a pseudo-Siamese point-tracking network and a feature-transfer network. The pose of the knee joints was estimated by tracking selected points on knee joints with the multi-view point-based registration network, paired DRRs, and fluoroscopy. A feature extractor was trained by the feature-learning network with pairs of DRRs and fluoroscopic images. To overcome the limited number of authentic fluoroscopic images, we trained the multi-view point-based registration network with DRRs and pre-trained the feature-learning network on ImageNet.

The remainder of this paper is organized as follows. Section 2 reviews deep-learning-based 2D–3D registration and domain adaption. Section 3 presents the proposed learning-based 2D–3D registration problems. Section 4 presents the experiments and results, and Section 5 concludes the paper.

## 2. Related work

### 2.1. Learning-based strategy

To avoid the large computational costs of optimization-based registration, researchers have recently developed learning-based registration [24]. Considering the success of convolutional neural networks (CNNs), feature extraction from both DRRs and fluoroscopic images has been proposed. The pose of the rigid object was then estimated by a hierarchical regressor [33]. The CNN model improves the robustness of registration, but it is limited to objects with strong features, such as medical implants, and cannot perform the registration of native anatomic structures. Miao et al. [28] proposed a reinforcement learning network to register X-ray and CT images of the spine with a Markov decision process. Although they improved the method with a multi-agent system, the proposed method may still fail because it cannot converge during searching. Recently, several attempts have been made to register rigid objects with point correspondence networks [27,34,35], which showed good results in both efficiency and accuracy on anatomic structures. Their method avoids the costly and unreliable iterative pose searching and corrects the out-of-plane errors with multiple views.

### 2.2. Domain adaption

The discrepancy between synthetic data (i.e., DRRs) and authentic data (i.e., fluoroscopic images), also known as drift, is another challenge to learning-based registration strategies, in which training data and future data must be in the same feature space and have the same distribution [36]. Compared with building complicated models for DRR generation, domain adaption has emerged as a promising and relatively effortless strategy to account for the domain difference between different image sources [37], and it has been applied in many medical applications, such as X-ray segmentation [38] and multi-modal image registration [21,22,39]. For 2D–3D registration, Zheng et al. [21] proposed the integration of a pairwise domain adaptation module into a pre-trained CNN that performs rigid registration using a limited amount of training data. The network was trained on DRRs, and it performed well on synthetic data; therefore, the authentic features were transferred close to the synthetic features with domain adaption. However, existing methods are still inappropriate for natural joints, such as knees and hips. Therefore, a designed registration approach for natural joints that do not require numerous clinical X-ray images for training is needed.

## 3. Methods

The aim of 2D–3D registration is to estimate the six degrees of freedom (6DOF) pose of 3D volume data from pairs of 2D multi-view fluoroscopic images. In the following section, we begin with an overview of the tracking system and multi-view point-based 2D–3D registration (Section 3.1). Then, details of the two main components of our work are given in Section 3.2 and Section 3.3.

### 3.1. Multi-view point-based registration

#### 3.1.1. 2D–3D rigid registration with 6DOF

We consider the registration of each bone in the knee joint as a separate 2D–3D registration procedure. Pose reproduction of each bone is denoted as the 3D alignment of the CT volume data $V$ through

a transformation matrix $T_{4\times4}$, which is parameterized by six elements of translations and rotations $(x, y, z, \gamma, \alpha, \beta)$ using the Euler angle [40]. Transformation matrix $T_{4\times4}$ can be represented as a homogeneous $4 \times 4$ matrix, and pose $P$ can be derived as follows:

$$P = T_{4\times4} \cdot V = \begin{bmatrix} R_{3\times3}(\gamma, \alpha, \beta) & t \\ 0 & 1 \end{bmatrix} \cdot V \tag{1}$$

where $R_{3\times3}$ is a rotation matrix about three axes $R_{3\times3} = R_\gamma \cdot R_\alpha \cdot R_\beta$, $R_\gamma = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $R_\alpha = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}$, and $R_\beta = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}$, and $t$ is translation vector along the three axes $t = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$.

#### 3.1.2. 3D projection geometry of X-ray imaging

In the virtual DFIS, the four corners of each imaging plane and the location of the X-ray sources were used to establish the optical pinhole model during DRR generation (Fig. 1). After a polynomial-based distortion correction and spatial calibration of two-view fluoroscopy, DRRs were generated by the ray-casting algorithm [41] with segmented CT volume data using Amira software (ThermoFisher Scientific, USA). Combing the transformation matrix $T_{4\times4}$, the final DRR $I_{\mathrm{DRR}}$ can be computed as follows:

$$I_{\mathrm{DRR}} = \int \mu\left[T_{4\times4}^{-1} \cdot l(p, s)\right] \mathrm{d}s \tag{2}$$

where $l(p, s)$ is the ray $s$ connecting the X-ray source and image plane in the X-ray imaging model, and $p$ is a point of the ray. $\mu(\cdot)$ represents the attenuation coefficient at some point in the volume data.

#### 3.1.3. Registration by tracking multi-view points

Previous literature has reported single-view 2D–3D registration to be an ill-posed problem; therefore, two-view fluoroscopic images were used for registration to avoid out-of-plane errors [42]. Considering the excellent performance of the point-based registration method on anatomic structures [27,34,35], we measured the dynamic motion of knee joints by tracking a set of selected points on the surface model in DFIS (Fig. 2), and we denoted the selected points as $P_{\mathrm{bone}} = [p^1, p^2, p^3, ..., p^N]$. 2D projection of the selected points was tracked with a pseudo-Siamese multi-view point-based registration network (Section 3.2). After tracking the selected points from all the provided views, we reproduced the 3D locations of the set of points $P_{\mathrm{E}} = [p^1_{\mathrm{estimated}}, p^2_{\mathrm{estimated}}, p^3_{\mathrm{estimated}}, \cdots, p^N_{\mathrm{estimated}}]$ using triangulation [43]. To determine the final transformation matrix $T$, a Procrustes analysis [44] was used as follows:

$$T = \arg\min_T \|T \cdot P_{\mathrm{bone}} - P_{\mathrm{E}}\|_2 \tag{3}$$

The final pose of each bone was reproduced with transformation matrix $T$.

### 3.2. Pseudo-Siamese point tracking network

In the proposed method, we used a pseudo-Siamese network to track points from each view. The pseudo-Siamese network has two branches: One is a visual geometry group (VGG) network [45] for extracting features from DRRs, and the other is a feature-transfer network, which transfers authentic features to synthetic features (Section 3.3). The overall workflow is shown in Fig. 3. The input of the network was unpaired DRRs and fluoroscopic images, and the output was the tracked points of the fluoroscopic images. In the upper branch of the network (Fig. 3), the exported features $F_{\mathrm{DRR}}$ around each selected point have the size of $M \times N \times C$ when the width and height of the DRR are respectively $M$ and $N$, and $C$ is the number of feature channels. In the lower branch of the network, the features of fluoroscopic images $F_{\mathrm{fluoro}}$, were exported by the feature-transfer network without weight sharing. With the output of the extracted features $F_{\mathrm{DRR}}$ and $F_{\mathrm{fluoro}}$, a convolutional layer was applied to quantify the similarity between the two feature maps [27]. The similarity is denoted as

$$M = F_{\mathrm{fluoro}}(W \odot F_{\mathrm{DRR}}) \tag{4}$$

where $W$ is a learned weighting factor in finding better similarity for each selected point. The objective function to be minimized during the training is Euclidean loss (i.e., registration loss), defined as

$$\varphi = \frac{1}{N}\sum_1^N \|p^i_{\mathrm{fluoro}} - p^i_{\mathrm{drr}}\|_2^2 \tag{5}$$

where $p_{\mathrm{fluoro}}$ is the tracked 2D points and $p_{\mathrm{drr}}$ is the projected 2D points in DRR with known locations. With the tracked 2D points from different views, the 3D points were reconstructed using triangulation [43].

### 3.3. Feature transfer using domain adaption

For feature extraction of fluoroscopic images, we proposed a transfer-learning-based method to reduce the domain difference between synthetic images (e.g., the DRRs) and authentic X-ray images (e.g., the fluoroscopic images) (Fig. 4).

To close the gap between the two domains, we used a domain-adaption method. That is, additional coupled VGG net with cosine similarity was set during feature extraction of the fluoroscopic images to close the gap (Fig. 5). Pairs of DRRs and fluoroscopic

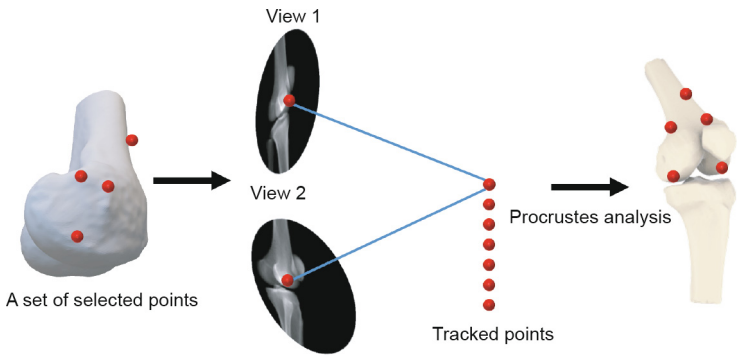


**Fig. 2.** The workflow of the multi-view point-based registration method. A set of points was selected on the bone surface, and their 2D projections were tracked from each view in the virtual DFIS to reconstruct their 3D positions. The final transformation matrix was determined by the reconstructed points using Procrustes analysis [44].

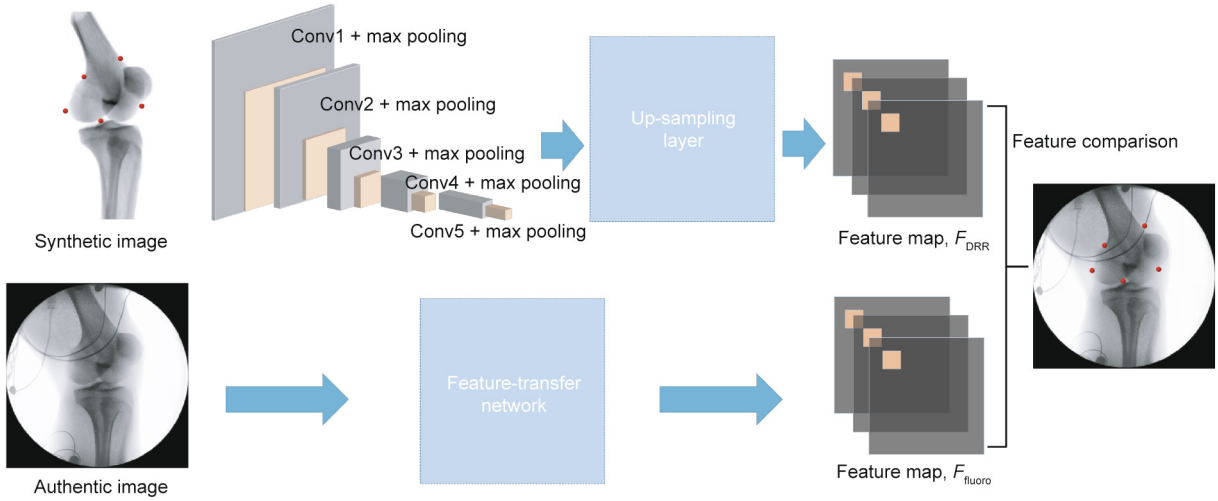


**Fig. 3.** The framework of the point-tracking network. Pairs of DRRs and fluoroscopic images were imported to the network, and their features were extracted by a VGG and a feature-transfer network, respectively. The selected points were tracked on fluoroscopic images by searching the most similar feature patch around the selected points in DRRs. Conv: convolution layers.

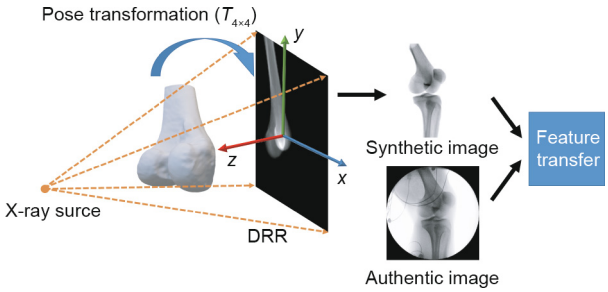


**Fig. 4.** Feature-transfer network with the paired synthetic image and authentic image. Synthetic images (i.e., DRRs) were generated at the pose after manual registration.

images, which share the same locations of volume data using a model-based manual registration method [9], were used for training. We used cosine similarity as the cost function to measure the gap between the two domains. For the tracking problem, the cosine similarity can be stated as

$$L\left(F^{X},F^{D}\right)=1-\frac{1}{M}\frac{1}{N}\sum_{1}^{M}\sum_{1}^{N}\frac{\langle F^{X},F^{D}\rangle}{\|F^{X}\|\|F^{D}\|}, \tag{6}$$

where $\|\cdot\|$ denotes $L_2$-norm and $\langle\cdot\rangle$ denotes dot product, and $F^{X}$ and $F^{D}$ are the feature maps. To improve the performance of feature transfer, we optimized the proposed method with weights pre-trained on ImageNet.

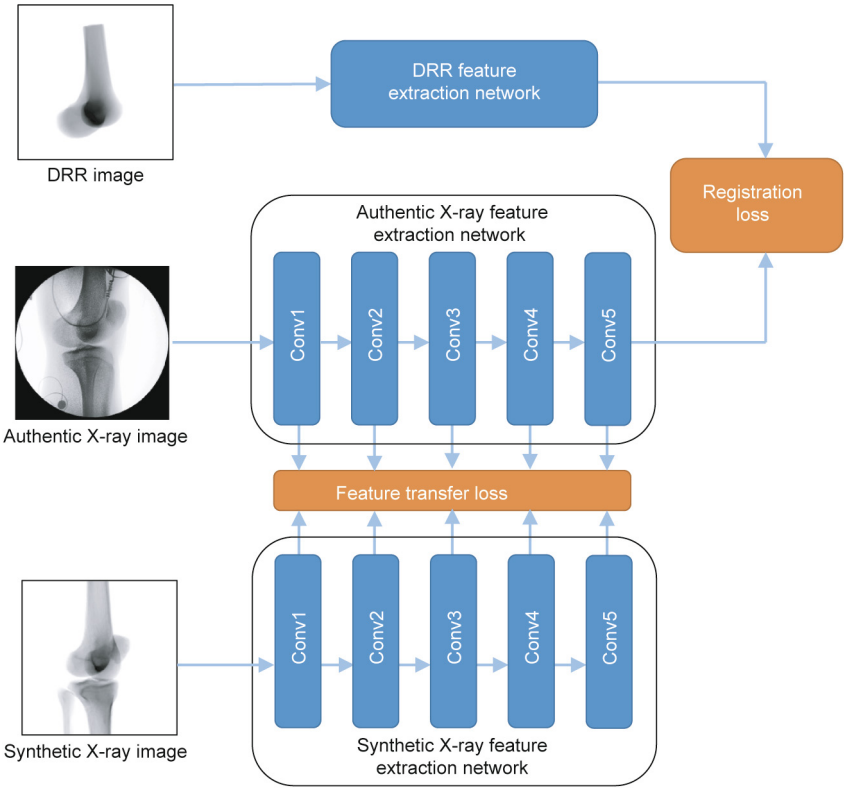


**Fig. 5.** The architecture of synthetic X-ray image feature extraction.

## 4. Experiments and results

### 4.1. Dataset

In this institutional-review-board-approved study, we collected CT images of three subjects' knees, and all subjects performed two or three motions that were captured by a bi-plane fluoroscopy system (BV Pulsera, Philips, the Netherlands) with a frame rate of 30 frames per second. CT scans (SOMATOM Definition AS; Siemens, Germany) of each knee, ranging from approximately 30 cm proximal and distal to the knee joint line (thickness, 0.6 mm; resolution 512 × 512 pixels), were obtained. The size of the fluoroscopic images was 1024 × 1024 pixels with a pixel spacing of 0.28 mm. Geometric parameters of the bi-plane fluoroscopy imaging model, such as polynomial distortion correction parameters [46] and the locations of the X-ray source and detector plane, were used to establish a virtual DFIS, in which poses of each bone were reproduced manually [47]. In this study, 143 pairs of matched fluoroscopic images were used (Fig. 6), of which 91 pairs of matched images were used for training the feature-transfer network of fluoroscopic images and the point tracking network, and the remaining images were used as the testing set. Additionally, a three-fold validation was performed in the study. To evaluate the 2D–3D registration algorithm, a widely used 3D error measurement (i.e., the target registration error (TRE)) was applied [48]. We computed the mean TRE (mTRE) to determine the 3D error. The average distance between the selected points defines the mTRE.

$$\mathrm{mTRE}(P_{\mathrm{E}}, P_{\mathrm{bone}}) = \frac{1}{k}\sum_{1}^{M}\|P_{\mathrm{E}} - P_{\mathrm{bone}}\| \tag{7}$$

where $P_{\mathrm{bone}}$ denotes the selected points and $P_{\mathrm{E}}$ denotes the estimated points. The success rate was defined as the percentage of all the test cases with an mTRE of less than 10 mm.

### 4.2. Loss selection in cross-domain feature extraction analysis

We defined a cosine similarity as the loss function in the feature extraction on the authentic X-ray images. We also used the mean squared error as the loss function [22] to find a better loss function. The position of the loss function may also affect the final performance of the feature extraction layer. Thus, we first compared the effects of loss functions located at different convolution layers. To obtain the best performance of the cross-domain feature from the real fluoroscopic images, we put the defined loss function between the pairs of conv2 layers, conv3 layers, conv4 layers, and conv5 layers. In our data (Fig. 7), we preferred the cosine similarity as the loss function because it has better performance regarding the final registration result of the entire knee joint. Cosine similarity showed the best performance between conv5 layers (see details in Appendix A, Table S1).

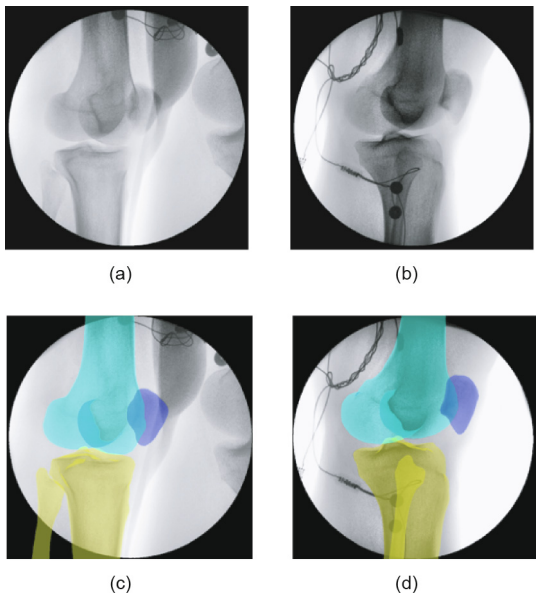


**Fig. 6.** Paired raw fluoroscopic images and the corresponding images after manual matching. The raw fluoroscopic images are (a) and (b), in which additional noise (wearable electromyography sensors) can be found on the surface of the lower limb. As described in the previous study [6], manual registration was performed until the projections of the surface bone model matched the outlines of the fluoroscopic images, and the matched results are shown in (c) and (d). Reproduced from Ref. [6] with permission of Elsevier Ltd., © 2011.

### 4.3. With or without transfer training network analysis

To test the effects of the proposed feature-based transfer learning method, we compared this method with the Siamese registration network (i.e., POINT$^2$ network) [27]. Moreover, as a widely used transfer learning tool, fine-tuning, was also compared in the current study to find a better way to reduce the differences

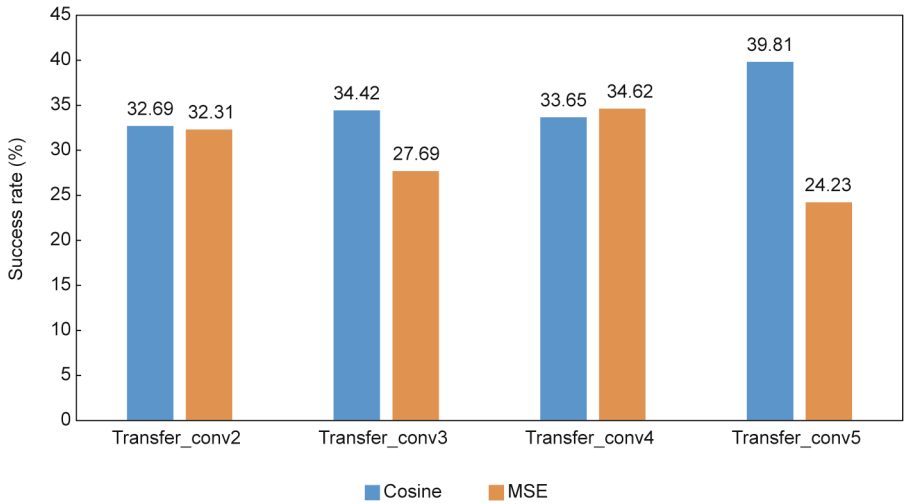


**Fig. 7.** The success rate using cosine similarity and mean squared error (MSE) at different convolutional layers.

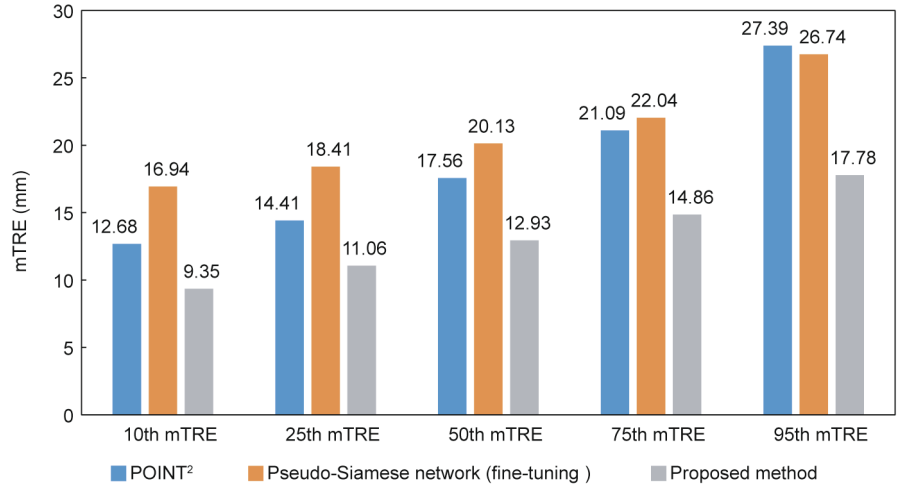


**Fig. 8.** Mean target registration error with different registration networks.

between the fluoroscopic images and DRRs. The weights of the proposed method were pre-trained on the ImageNet database. The average performance of 10 tests for each method was used as the final performance. The mTRE results are reported in terms of the 10th, 25th, 50th, 75th, and 95th percentiles to demonstrate the robustness of the compared methods. The proposed feature-based transfer learning method had a significantly better performance than the Siamese registration network (Fig. 8), and it also performed better than fine-tuning, with a success accuracy rate of almost zero (Table S2 in Appendix A).

### 4.4. Three-fold cross-validation

We used three-fold cross-validation in this study and compared the proposed pseudo-Siamese registration network with and without transfer learning. Therefore, two of the three subjects were used for training the system, and the last subject was used to validate the system. This approach was iterated ten times by shifting the test subjects randomly. The performances (mTRE) were evaluated in each iteration. Finally, the performances recorded in all ten iterations were averaged to obtain a final mTRE. The mTRE results are reported in terms of the 10th, 25th, 50th, 75th, and 95th percentiles (Table 1). The final three-fold cross-validation showed that the proposed method also had a better performance with feature transfer.

## 5. Conclusions

To overcome limited numbers of real fluoroscopic images in learning-based 2D–3D rigid registration via DRRs, we proposed a pseudo-Siamese multi-view point-based registration framework. The proposed method can decrease the demand for real X-ray images. With the ability to transfer authentic features to synthetic features, the proposed method has better performance than the fine-tuning pseudo-Siamese network. This study also estimated the POINT$^2$ network with and without transfer learning. The results showed that the proposed pseudo-Siamese network has a better success rate and accuracy than the Siamese point-tracking network. With a small amount of training data, the proposed method can work as an initialization step for the optimization-based registration method to improve accuracy. However, there are several limitations to the current work. First, because our method is designed for at least two fluoroscopic views, multi-view data were required to reconstruct the knee poses; otherwise, out-of-plane translation and rotation error would be large because of the physical imaging model. Second, the proposed method cannot reach a sub-millimeter accuracy compared with an optimization-based strategy. Like other learning-based strategies, our proposed method did not provide good accuracy but would be much faster than the optimization-based method, because no iterative step was needed during matching. In clinical orthopedic practice, accurate joint kinematics is essential for the determination of a

**Table 1**
Three-fold cross-validation with and without transfer learning.

| Term | Registration network with transfer learning | | | | Registration network without transfer learning | | | |
|---|---|---|---|---|---|---|---|---|
| | Femur | Patella | Tibia | Joint[a] | Femur | Patella | Tibia | Joint[a] |
| 10th mTRE | 6.45 | 6.68 | 5.52 | 8.16 | 5.71 | 7.49 | 5.84 | 7.47 |
| 25th mTRE | 8.86 | 8.27 | 7.07 | 9.28 | 10.44 | 8.50 | 7.66 | 9.61 |
| 50th mTRE | 12.17 | 10.73 | 9.09 | 11.08 | 16.74 | 9.61 | 10.53 | 12.70 |
| 75th mTRE | 17.52 | 15.23 | 11.58 | 13.58 | 21.20 | 11.34 | 13.43 | 14.57 |
| 95th mTRE | 26.68 | 27.61 | 15.99 | 20.49 | 27.64 | 16.16 | 17.55 | 18.37 |
| Success rate | 0.36 | 0.42 | 0.60 | 0.35 | 0.24 | 0.57 | 0.45 | 0.28 |

All values are in millimeters.
[a] Joint means the final registration result of the whole joint.

rehabilitation scheme [5], surgical planning [1], and functional evaluation [47]. The proposed method alone is inappropriate for *in-vivo* joint kinematics measurement. Therefore, a combination of our method with an optimization-based strategy would be a viable solution.

## Acknowledgements

This project was sponsored by the National Natural Science Foundation of China (31771017, 31972924, and 81873997), the Science and Technology Commission of Shanghai Municipality (16441908700), the Innovation Research Plan supported by Shanghai Municipal Education Commission (ZXWF082101), the National Key R&D Program of China (2017YFC0110700, 2018YFF0300504 and 2019YFC0120600), the Natural Science Foundation of Shanghai (18ZR1428600), and the Interdisciplinary Program of Shanghai Jiao Tong University (ZH2018QNA06 and YG2017MS09).

## Compliance with ethics guidelines

Cong Wang, Shuaining Xie, Kang Li, Chongyang Wang, Xudong Liu, Liang Zhao, and Tsung-Yuan Tsai declare that they have no conflict of interest or financial conflicts to disclose.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eng.2020.03.016.

## References

[1] Kernkamp WA, Wang C, Li C, Hu H, Van Arkel ERA, Nelissen RGHH, et al. The medial patellofemoral ligament is a dynamic and anisometric structure: an *in vivo* study on length changes and isometry. Am J Sports Med 2019;47:1645–53.
[2] Li JS, Tsai TY, Clancy MM, Li G, Lewis CL, Felson DT. Weight loss changed gait kinematics in individuals with obesity and knee pain. Gait Posture 2019;68:461–5.
[3] Li G, Li JS, Torriani M, Hosseini A. Short-term contact kinematic changes and longer-term biochemical changes in the cartilage after ACL reconstruction: a pilot study. Ann Biomed Eng 2018;46:1797–805.
[4] Tsai TY, Liow MHL, Li G, Arauz P, Peng Y, Klemt C, et al. Bi-cruciate retaining total knee arthroplasty does not restore native tibiofemoral articular contact kinematics during gait. J Orthop Res 2019;37(9):1929–37.
[5] Perriman A, Leahy E, Semciw AI. The effect of open- versus closed-kinetic-chain exercises on anterior tibial laxity, strength, and function following anterior cruciate ligament reconstruction: a systematic review and meta-analysis. J Orthop Sports Phys Ther 2018;48(7):552–66.
[6] Tsai TY, Lu TW, Kuo MY, Lin CC. Effects of soft tissue artifacts on the calculated kinematics and kinetics of the knee during stair-ascent. J Biomech 2011;44(6):1182–8.
[7] Lu TW, O'Connor JJ. Bone position estimation from skin marker co-ordinates using global optimisation with joint constraints. J Biomech 1999;32(2):129–34.
[8] Alexander EJ, Andriacchi TP. Correcting for deformation in skin-based marker systems. J Biomech 2001;34(3):355–61.
[9] Lucchetti L, Cappozzo A, Cappello A, Della CU. Skin movement artefact assessment and compensation in the estimation of knee-joint kinematics. J Biomech 1998;31(11):977–84.
[10] Leardini A, Chiari L, Della Croce U, Cappozzo A. Human movement analysis using stereophotogrammetry: part 3. Soft tissue artifact assessment and compensation. Gait Posture 2005;21(2):212–25.
[11] Borotikar B, Lempereur M, Lelievre M, Burdin V, Ben Salem D, Brochard S. Dynamic MRI to quantify musculoskeletal motion: a systematic review of concurrent validity and reliability, and perspectives for evaluation of musculoskeletal disorders. PLoS ONE 2017;12(12):e0189587.
[12] Tanaka MJ, Elias JJ, Williams AA, Demehri S, Cosgarea AJ. Characterization of patellar maltracking using dynamic kinematic CT imaging in patients with patellar instability. Knee Surg Sports Traumatol Arthrosc 2016;24:3634–41.
[13] Patel VV, Hall K, Ries M, Lotz J, Ozhinsky E, Lindsey C, et al. A three-dimensional MRI analysis of knee kinematics. J Orthop Res 2004;22(2):283–92.
[14] Muhle C, Brossmann J, Heller M. Kinematic CT and MR imaging of the patellofemoral joint. Eur Radiol 1999;9:508–18.
[15] Li G, Wuerz TH, DeFrate LE. Feasibility of using orthogonal fluoroscopic images to measure *in vivo* joint kinematics. J Biomech Eng 2004;126(2):314–8.
[16] Acker S, Li R, Murray H, St JP, Banks S, Mu S, et al. Accuracy of single-plane fluoroscopy in determining relative position and orientation of total knee replacement components. J Biomech 2011;44(4):784–7.
[17] Kim J, Li S, Pradhan D, Hammoud R, Chen Q, Yin FF, et al. Comparison of similarity measures for rigid-body CT/dual X-ray image registrations. Technol Cancer Res Treat 2007;6(4):337–46.
[18] Van der Bom IMJ, Klein S, Staring M, Homan R, Bartels LW, Pluim JPW. Evaluation of optimization methods for intensity-based 2D–3D registration in X-ray guided interventions. In: Proceedings of SPIE Medical Imaging; 2011 Feb 12–17; Orlando, FL, USA; 2011.
[19] Alam F, Rahman SU, Khusro S, Ullah S, Khalil A. Evaluation of medical image registration techniques based on nature and domain of the transformation. J Med Imaging Radiat Sci 2016;47(2):178–93.
[20] Ghafurian S, Hacihaliloglu I, Metaxas DN, Tan V, Li K. A computationally efficient 3D/2D registration method based on image gradient direction probability density function. Neurocomputing 2017;229:100–8.
[21] Zheng J, Miao S, Wang ZJ, Liao R. Pairwise domain adaptation module for CNN-based 2-D/3-D registration. J Med Imaging 2018;5(2):21204.
[22] Zheng J, Miao S, Liao R. Learning CNNs with pairwise domain adaption for real-time 6DoF ultrasound transducer detection and tracking from X-ray images. In: Proceedings of the 20th International Conference on Medical Image Computing and Computer-Assisted Intervention; 2017 Sep 11–13; Quebec City, QC, Canada; 2017.
[23] Markelj P, Tomaževič D, Likar B, Pernuš F. A review of 3D/2D registration methods for image guided interventions. Med Image Anal 2012;16(3):642–61.
[24] Haskins G, Kruger U, Yan P. Deep learning in medical image registration: a survey. Mach Vis Appl 2020;31(1):8.
[25] Duan Q, Wang G, Wang R, Fu C, Li X, Gong M, et al. SenseCare: a research platform for medical image informatics and interactive 3D visualization 2020. arXiv:2004.07031v1.
[26] Toth D, Miao S, Kurzendorfer T, Rinaldi CA, Liao R, Mansi T, et al. 3D/2D model-to-image registration by imitation learning for cardiac procedures. Int J Comput Assisted Radiol Surg 2018;13:1141–9.
[27] Liao H, Lin WA, Zhang J, Zhang J, Luo J, Zhou SK. Multiview 2D/3D rigid registration via a point-of-interest network for tracking and triangulation. In: Proceedings of the 2019 IEEE Conference on Computer Vision Pattern Recognition; 2019 Jun 15–20; Long Beach, CA, USA; 2019. p. 12638–47.
[28] Miao S, Piat S, Fischer P, Tuysuzoglu A, Mewes P, Mansi T, et al. Dilated FCN for multi-agent 2D/3D medical image registration. In: Proceedings of the 32nd AAAI Conference on Artificial Intelligence; 2018 Feb 2–7; New Orleans, LA, USA; 2018.
[29] Wallace WA, Johnson F. Detection and correction of geometrical distortion in X-ray fluoroscopic images. J Biomech 1981;14(2):123–5.
[30] Cerciello T, Bifulco P, Cesarelli M, Fratini A. A comparison of denoising methods for X-ray fluoroscopic images. Biomed Signal Process Control 2012;7(6):550–9.
[31] Moore CS, Liney GP, Beavis AW, Saunderson JR. A method to produce and validate a digitally reconstructed radiograph-based computer simulation for optimisation of chest radiographs acquired with a computed radiography imaging system. Br J Radiol 2011;84(1006):890–902.
[32] Moore CS, Wood TJ, Saunderson JR, Beavis AW. A method to incorporate the effect of beam quality on image noise in a digitally reconstructed radiograph (DRR) based computer simulation for optimisation of digital radiography. Phys Med Biol 2017;62(18):7379–93.
[33] Miao S, Wang ZJ, Zheng Y, Liao R. Real-time 2D/3D registration via CNN regression. In: Proceedings of the 2016 IEEE 13th International Symposium Biomedical Imaging; 2016 Apr 13–16; Prague, Czech Republic; 2016. p. 1430–4.
[34] Schaffert R, Wang J, Fischer P, Borsdorf A, Maier A. Metric-driven learning of correspondence weighting for 2-D/3-D image registration. In: Proceedings of the German Conference Pattern Recognition; 2018 Oct 9–12; Stuttgart, Germany; 2018. p. 140–52.
[35] Pan J, Min Z, Zhang A, Ma H, Meng MQH. Multiview global 2D–3D registration based on branch and bound algorithm. In: Proceedings of the 2019 IEEE International Conference on Robotics and Biomimetics; 2019 Dec 6–8; Dali, China; 2019. p. 3082–7.
[36] Weiss K, Khoshgoftaar TM, Wang D. A survey of transfer learning. J Big Data 2016;3:9.
[37] Csurka G. A comprehensive survey on domain adaptation for visual applications. In: Csurka G, editor. Domain adaptation in computer vision applications. Cham: Springer International Publishing AG; 2017. p. 1–35.
[38] Zhang Y, Miao S, Mansi T, Liao R. Task driven generative modeling for unsupervised domain adaptation: application to X-ray image segmentation. In: Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention; 2018 Sep 16–20; Granada, Spain; 2018. p. 599–607.
[39] Hu J, Sun S, Yang X, Zhou S, Wang X, Fu Y, et al. Towards accurate and robust multi-modal medical image registration using contrastive metric learning. IEEE Access 2019;7:132816–27.
[40] Otake Y, Armand M, Armiger RS, Kutzer MD, Basafa E, Kazanzides P, et al. Intraoperative image-based multiview 2D/3D registration for image-guided orthopaedic surgery: incorporation of fiducial-based C-arm tracking and GPU-acceleration. IEEE Trans Med Imaging 2012;31(4):948–62.
[41] Li T, Zhu H. Research on digitally reconstructed radiograph algorithm improvement based on computed tomography image. In: Proceedings of the 2018 11th International Congress on Image Signal Processing, Biomedical Engineering Informatics; 2018 Oct 13–15; Beijing, China; 2018. p. 1–5.

[42] Fischer B, Modersitzki J. Ill-posed medicine—an introduction to image registration. Inverse Probl 2008;24:034008.
[43] Hartley R, Zisserman A. Multiple view geometry in computer vision. Cambridge: Cambridge University Press; 2003.
[44] Seber GAF. Multivariate observations. New York: John Wiley & Sons; 1983.
[45] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014. arXiv:1409.1556.
[46] Baltzopoulos V. A videofluoroscopy method for optical distortion correction and measurement of knee-joint kinematics. Clin Biomech 1995;10(2):85–92.
[47] Tsai TY, Dimitriou D, Li G, Kwon YM. Does total hip arthroplasty restore native hip anatomy? Three-dimensional reconstruction analysis. Int Orthop 2014;38:1577–83.
[48] Van de Kraats EB, Penney GP, Tomazevic D, Van Walsum T, Niessen WJ. Standardized evaluation methodology for 2-D–3-D registration. IEEE Trans Med Imaging 2005;24(9):1177–89.