

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/329562190>

Learning Reasoning–Decision Networks for Robust Face Alignment

Article in IEEE Transactions on Pattern Analysis and Machine Intelligence · December 2018

DOI: 10.1109/TPAMI.2018.2885298

CITATIONS

0

READS

86

5 authors, including:



Hao Liu

Ningxia University

8 PUBLICATIONS 71 CITATIONS

[SEE PROFILE](#)



Jie Zhou

Linköping University

447 PUBLICATIONS 7,184 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Left Atrial Appendage Segmentation [View project](#)



Deep Learning [View project](#)

Learning Reasoning-Decision Networks for Robust Face Alignment

Hao Liu, Jiwen Lu, *Senior Member, IEEE*, Minghao Guo, Suping Wu, Jie Zhou, *Senior Member, IEEE*

Abstract—In this paper, we propose an end-to-end reasoning-decision networks (RDN) approach for robust face alignment via policy gradient. Unlike the conventional coarse-to-fine approaches which likely lead to bias prediction due to poor initialization, our approach aims to learn a policy by leveraging raw pixels to reason a subset of shape candidates, sequentially making plausible decisions to remove outliers for robust initialization. To achieve this, we formulate face alignment as a Markov decision process by defining an agent, which typically interacts with a trajectory of states, actions, state transitions and rewards. The agent seeks an optimal shape searching policy over the whole shape space by maximizing a discounted sum of the received values. To further improve the alignment performance, we develop an LSTM-based value function to evaluate the shape quality. During the training procedure, we adjust the gradient of our value function in directions of the policy gradient. This prevents our training goal from being trapped into local optima entangled by both the pose deformations and appearance variations especially in unconstrained environments. Experimental results show that our proposed RDN consistently outperforms most state-of-the-art approaches on four widely-evaluated challenging datasets.

Index Terms—Face alignment, deep neural networks, deep reinforcement learning, policy gradient.

1 INTRODUCTION

Face alignment (*a.k.a.* facial landmark localization) aims to localize multiple facial landmarks for a given facial image, which plays a significant pre-processing role in diverse facial analysis tasks such as face identification and verification [23], [24], [66], head pose estimation [76], [87] and facial attribute analysis [4], [34]. While extensive efforts have been devoted to face alignment, the performance still remains unsatisfied in practice especially when face images were captured in unconstrained environments. This is mainly because these in-the-wild face images usually undergo significant variations due to large poses, diverse expressions and partial occlusions.

In the context of variant face alignment approaches [9], [64], [78], [85], cascade regression has been one of the state-of-the-arts. Methods in this category typically seek linear feature-to-shape mappings to refine the shape incrementally, starting with an initialization, *e.g.*, the statistically averaged shape. However, one major issue in these methods is that these linear mappings are too weak to exploit the complex and nonlinear relationship between image pixels and shape variations in unconstrained environments. To address this nonlinear issue, deep learning has been leveraged in face alignment recently [28], [42], [64], [68], [83], which aims to exploit the nonlinear mappings by a cascade of neural networks and achieve superior performance. Nevertheless, these methods are sensitive to low-quality initializations. The underlying reason

is that for the cascade executions, the latter stage proceeds to correct the misalignment errors, which intensely depends on the outputting shapes of previous stages. Consequently, the discrepancy between the entangled-initialized shape and the target shape cumulatively harms the alignment accuracy. Fig. 1(a) illustrates the failure case mainly caused by the low-quality initialization. Besides seeking discriminative feature-to-shape mappings, a well-defined alignment error function has been proven to be effective on promoting the alignment accuracy in recent studies [14], [26], [27], [33], [48], [74], [84]. For example, Jourabloo *et al.* [26], [27] introduced 3D geometric information and auxiliary facial attributes in their objective functions, so that more structural cues are exploited for model learning. The heat map-based methods such as [33], [74] enhance the point-wise measurement to regressing one 2D heat map for each landmark, so that these maps integrate their training objectives with an additional 2D spatial constraint to improve their performance. Nevertheless, both the point-wise and heat map-wise measurements only exploit partial 2D landmark locations, and ignore the complementary information of facial appearance variations and different landmark partitions for plausible shape decisions [5]. Moreover, these measurement functions are hand-designed which may be trapped into local optima, since they cannot provide the optimal supervision signals for efficient training. To fully exploit the global cues of shape decisions, a generic solution is to jointly optimize both procedures of learning an optimal objective function and seeking a cascade of plausible image-to-shape mappings, which can further improve the alignment accuracy.

To address the aforementioned challenges, we propose an end-to-end reasoning-decision networks (RDN) approach for robust face alignment especially in unconstrained environments. The basic goal of our RDN aims to reason a sequence of efficient descent directions for shape updates, in parallel to learning a globally-optimized shape evaluation function to accurately evaluate the shape quality. To achieve this, we formulate face alignment as a Markov decision process (MDP) by defining the agent to

• *Hao Liu is with the School of Information Engineering, Ningxia University, Yinchuan, 750021, China. He is also with the Department of Automation, Tsinghua University, Beijing, 100084, China. Suping Wu is with the School of Information Engineering, Ningxia University, Yinchuan, 750021, China. E-mail: liuhao@nxu.edu.cn; pswuu@nxu.edu.cn.*

• *Jiwen Lu, Minghao Guo and Jie Zhou are with the Department of Automation, Tsinghua University, State Key Lab of Intelligent Technologies and Systems, and Tsinghua National Laboratory for Information Science and Technology (TNList), Beijing, 100084, China. E-mail: lujiwen@tsinghua.edu.cn; gmh14@mails.tsinghua.edu.cn; jzhou@tsinghua.edu.cn.*

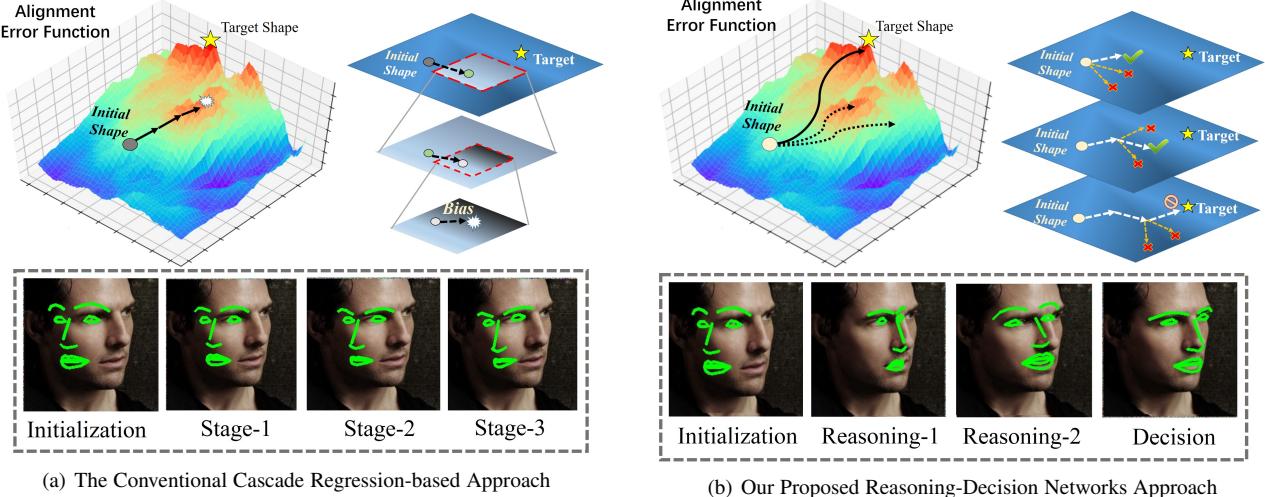


Fig. 1. Our proposed reasoning-decision networks framework versus the conventional regression-based approach. For the given testing face (near-frontal) with a poor initialization, the cascaded regression-based approach learns to correct misalignment errors intensely based on initialized shape and may cause bias prediction. Different from those approaches, our RDN architecture reasons a plausible shape searching policy over the whole shape space, which maximizes the cumulated quality values received by the learned shape evaluation function. Based on the obtained shape quality values, we successfully achieve to refine low-quality initialization for robust face alignment. Our approach will automatically terminates after iterations until reaching the alignment accuracy threshold.

interact with a trajectory of states, actions and rewards in the MDP. As illustrated in Fig. 1(b), our RDN agent seeks a plausible shape searching policy over the continuous shape space, which accumulates greater shape quality values received by the self-learned shape evaluation function. Our policy proceeds to reason a subset of shape candidates, sequentially making decisions on them to choose the highest-scoring candidate for robust initialization. Furthermore, we design a reward function to compute a scalar reward value. Then we utilize this value as the weakly-supervised signal to automatically learn the alignment evaluation function via reinforcement learning. More specifically, our RDN architecture incorporates with the shape-reasoning and shape-decision modules under the paradigm of deep neural networks. Based on the observed appearance information, the shape-reasoning module makes a shape inference with the convolutional neural networks (CNN) and maps it to a subset of neighbouring shapes as the candidates via K-NN shape searching constraint. Accordingly, our shape-decision module leverages individual facial parts to evaluate the shape quality of these candidates via long-short term memory (LSTM) and then selects the highest ranked one for the next-iteration initialization. During the model training procedure, we refine the policy parameters according to the gradient of the alignment evaluation function via policy gradient. Fig. 2 details the flowchart of our proposed RDN. To show the effectiveness of the proposed RDN, we conduct experiments on four challenging benchmarks and the experimental results show the effectiveness of our RDN compared with the state-of-the-arts.

The contributions of our work are summarized as follows:

- 1) Our proposed approach globally optimizes both procedures of learning the alignment error function and inferring descent directions via deep reinforcement learning. As a result, we achieve a global-optimized alignment error function to accurately evaluate the shape quality only by a scalar reward. Moreover, our model reasons a sequence of reasonable and efficient descent directions by receiving the greater cumulative reward. To the best of our knowledge,

we are the first work to apply reinforcement learning, which addresses face alignment as seeking an optimal searching policy over the whole 2D shape space.

- 2) Benefiting from the discounted correlation in the Markov decision process, our RDN acts in reasonable decisions to refine future-iteration initializations by preserving the K-NN shape searching constraint. This definitely prevents our face alignment model from being trapped into local optima due to low-quality initialization during the inference stage.

2 RELATED WORK

In this section, we briefly review existing face alignment methods and the deep reinforcement learning techniques.

2.1 Face Alignment

Existing face alignment approaches can be roughly divided into three categories: discriminative model-based [3], [7], [9]–[11], [28], [70], [78], [83]–[85], cascaded regression-based [3], [9], [29], [53], [69], [78], [85], [86] and deep learning-based [7], [28], [41], [42], [44], [64], [68], [75], [81], [83]. Discriminative model-based methods estimate facial landmarks by maximizing the joint posterior probability over all landmarks for the given input image. However, these methods usually undergo time-consuming issue and their computation loads are heavy, which cannot fulfill the real-time requirement in practice. To address this limitation, cascaded regression-based methods have been proposed to refine the facial shape based on the given initial shape in a coarse-to-fine manner, which have achieved large improvements in terms of both the efficiency and the accuracy. For example, Xiong *et al.* [78] proposed a supervised learning method to approximate a sequence of descent directions in order to settle the quadric optimization, which aims to refine the facial shape based on cascade executions. Zhu *et al.* [85] developed a coarse-to-fine searching approach to narrow down the shape searching space, which demonstrates the

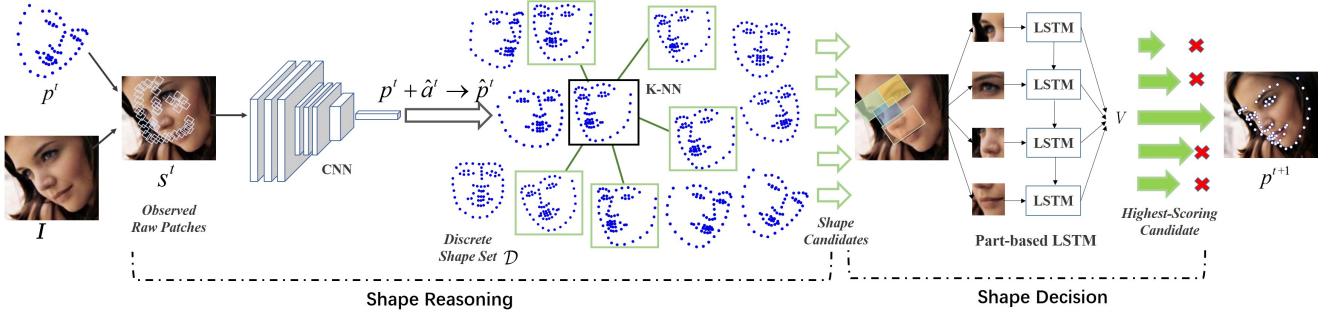


Fig. 2. The architecture of our proposed RDN. Our RDN starts with a given facial image I and an entangled-initialized shape p^0 . Taken the t -th iteration as an example, our shape reasoning module predicts an action \hat{a}^t to refine the shape as \hat{p}^t and produces a subset of shape candidates via K -NN shape searching constraint. Our shape decision module evaluates the quality value V for each shape candidate and then chooses the highest scored one as the future initialization p^{t+1} . It should be noted that we regard the immediate shape \hat{p}^t as final prediction, instead of these outputting shape set produced by K -NN searching.

outstanding performance on the standard benchmarking datasets. Nevertheless, the employed linear regressors are not powerful enough to exploit the complex and nonlinear relationship between the face data and facial shapes.

These years have witnessed that deep neural networks [21], [31], [49] have been applied to address the nonlinear issue in face alignment [7], [28], [33], [41], [44], [64], [68], [75], [77], [83], which learns a set of nonlinear hierarchical mappings between facial appearances and shape deformations in an end-to-end manner. For example, Trigeorgis *et al.* [68] proposed an end-to-end deep architecture to memorize the descent directions across consecutive refinement stages. Liu *et al.* [41] presented a deep convolutional neural networks architecture, which aims at exploiting the structural information of neighbouring landmarks and meanwhile sharing the parameters across stages. Although these methods have achieved the promising performance, one major issue is that the cascaded execution proceeds intensive dependency across consecutive stages, cumulatively causing the degraded performance under the challenging circumstances due to large poses, diverse expressions and severe occlusions. Moreover, recent deep models [14], [27], [33], [48], [74], [84] have demonstrated that superior-designed alignment error functions help to improve the performance. However, these hand-crafted functions are not optimal for shape evaluation and efficient training, because they cannot exploit the full supervision signals incorporating with facial appearance variations and semantic landmark partitions [5], [85]. To address these problems, our proposed method jointly optimizes both procedures of finding a set of plausible descent directions and learning a globally-optimal alignment error function under a unified deep neural networks architecture.

2.2 Deep Reinforcement Learning

Recent trends [22], [37], [45], [71], [72] have been seen that reinforcement learning is gradually introduced to many computer vision applications [25], [32], [43], [51], [52], [54], [63], [65], [67], [82], which teaches a policy to select actions sequentially by obtaining a greater feedback value each time. Policy gradient [73], derived as one of popular approaches in the reinforcement learning family, directly learns a policy function by maximizing the cumulated reward via standard gradient descent method [2]. Inspired by the success of deep learning [36], latest works [15], [22], [38], [61], [62] leverage deep neural networks to substitute the policy

function to further improve the performance. For example, Silver *et al.* [62] proposed a deterministic policy gradient algorithm to improve the efficiency of policy gradient over the high-dimension action space. Inspired by this, Lillicrap *et al.* [38] showed the deterministic policy gradient can operate on the continuous action space with deep neural networks. Although some efforts have been made on face recognition and detection [50]–[52], self-reasoning the descent directions by globally utilizing shape decisions for modern face alignment has not been visited yet.

In this paper, we consider face alignment as an action-decision process inspired by the REINFORCE [73] algorithm, where the descent directions for reasonable shape refinements are refined in directions of the gradients for the alignment error function. Specifically, we formulate face alignment as a Markov decision process and define an agent, which acts in shape reasoning to make decisions on the highest-scored shape quality value received from the value function. Moreover, we carefully design a reward function to compute a scalar value as the supervision signal. Our model requires this weak signal to learn the alignment error function by deep reinforcement learning. In parallel, we reason a sequence of descent directions with the K -NN searching constraint to remove the outliers for robust initialization. In this way, an optimal shape searching path is self-reasoned over the global 2D shape space by maximizing the greater shape quality values across iterations. During the training procedure, we adjust the model parameters in directions which seek the greater cumulative reward via deterministic policy gradient [62]. Consequently, our model achieves to jointly optimize both phases of learning an optimal function and finding efficient descents for plausible facial landmark localization.

3 APPROACH

Our basic idea of this paper is to globally optimize an alignment error function, in parallel to seeking a sequence of efficient descent directions for robust shape inference. As demonstrated in Fig. 2, we formulate face alignment problem as a *Markov decision process* (MDP) and further propose a reasoning-decision networks (RDN) architecture. Our MDP agent typically reasons a subset of shape candidates over the global shape space at each iteration, in parallel to making plausible decisions for future-iteration initialization. Before then, we present the MDP formulation with respect to reinforcement learning.

3.1 MDP Formulation

Given a facial image \mathbf{I} , we let $\mathbf{p} = [p_1, p_2, \dots, p_L] \in \mathcal{P} \in \mathbb{R}^{2 \times L}$ denote the shape vector with L points, where p_i represents the horizontal and vertical coordinates for the i -th landmark. Let the vector $\mathbf{p}^* = [p_1^*, \dots, p_L^*]$ denote all the points for the targeting groundtruth shape. In this work, our architecture defines an agent to interact with the MDP by an action space, a state space, a state transition function, a reward function and a policy to select actions.

Action: The MDP *action* specifies the shape residual $\mathbf{a} \in \mathbb{R}^{2 \times L}$ over a continuous space, which provides a deformable offset to adjust all landmarks to the refined positions.

State: The MDP *state* is defined by a set of observed raw patches $\mathbf{s} = o(\mathbf{I}, \mathbf{p}) \in \mathcal{S} \in \mathbb{R}^{d \times d \times L}$, which are locally cropped around each facial landmark from the facial image in a shape-indexed manner [9], [53], [68], [78], where d denotes the length of each local patch. The initial state starts from \mathbf{p}^0 and the desired state requires to accurately capture the semantic facial parts such as eyes, eyebrow, nose, mouth and facial contour.

State Transition: The MDP *state transition* is composed by both transitions of shapes and features. Having received a new action \mathbf{a}^t at the t -th iteration, the shape is adjusted by the shape transition $\mathbf{p}^{t+1} = \mathbf{p}^t + \mathbf{a}^t$, while the observed patches is shifted by the feature transition as $\mathbf{s}^{t+1} = o(\mathbf{I}, \mathbf{p}^{t+1})$.

Reward: The MDP *reward* function is designed to measure the misalignment descent by the normalized point-to-point distance [55], which is defined as follows:

$$r(\mathbf{s}^t, \mathbf{a}^t) = \begin{cases} 1, & \text{if } m^t - m^{t+1} \geq \epsilon, \\ -1, & \text{if } m^t - m^{t+1} < 0, \\ m^t - m^{t+1}, & \text{otherwise,} \end{cases} \quad (1)$$

where ϵ denotes an empirically thresholding value, the $m^t = \frac{\sum_{i=1}^L \|p_i^t - p_i^*\|}{L \cdot \zeta}$ at the t -th iteration, $\|\cdot\|$ specifies the ℓ_2 norm, and ζ denotes the normalizing factor, e.g., the inter-pupil distance [9] or the square root of the face size [26], [86], respectively.

Policy: Our MDP policy denoted by π aims to reason a plausible shape searching path globally over the continuous shape space, which maximizes the expected shape quality values over total iterations. It is difficult to leverage the stochastic policy to reason continuous actions, because finding an optimal policy requires more samples and proceeds costly, especially if the dimension of the action space is large [62]. To circumvent this problem, we maintain a deterministic and differential policy function $\mathbf{a} = f_{\theta_\pi}(\mathbf{s})$ parameterized by θ_π under the CNN architecture, which is leveraged to exploit the nonlinear mapping between pairs of states and actions.

Objective Function: Starting from the distribution density $\varphi_1(\mathbf{s})$ of the initial state, we define the probability distribution $\varphi(\mathbf{s} \rightarrow \mathbf{s}', t, \pi)$ at the state \mathbf{s}' after transitioning across t iterations from the state \mathbf{s} . The discounted state distribution is formulated as $\rho_\pi(\mathbf{s}') := \int_{\mathcal{S}} \sum_{t=1}^{\infty} \gamma^{t-1} \varphi_1(\mathbf{s}) \varphi(\mathbf{s} \rightarrow \mathbf{s}', t, \pi) d\mathbf{s}$, where γ denotes a discounted factor valued in [0,1]. Therefore, the performance objective of our MDP agent can be written as the following expectation form:

$$\begin{aligned} J(\theta_\pi) &= \int_{\mathcal{S}} \rho_\pi(\mathbf{s}) r(\mathbf{s}, f_{\theta_\pi}(\mathbf{s})) d\mathbf{s} \\ &= \mathbb{E}_{\mathbf{s} \sim \rho_\pi} [r(\mathbf{s}, f_{\theta_\pi}(\mathbf{s}))], \end{aligned} \quad (2)$$

where $\mathbb{E}_{\mathbf{s} \sim \rho_\pi} [r(\mathbf{s}, f_{\theta_\pi}(\mathbf{s}))]$ denotes the expected value with respect to the discounted state distribution $\rho_\pi(\mathbf{s})$, and $f_{\theta_\pi}(\cdot)$

deterministically specifies our policy. Due to the benefit of the deterministic policy function, the expectation objective of learning policy integrates on only the state space, exploiting efficient exploration during the training process [38], [62].

3.2 Reasoning-Decision Networks

The main goal of our agent aims to obtain an optimal policy π which maximizes the cumulative and discounted reward across iterations. In this section, we allow the proposed RDN to approximate both the shape inference policy function and the shape evaluation function over the large state space and the continuous action space. To this end, our RDN contains the shape-reasoning module to produce a subset of shape candidates based on the observed states with CNN and the shape-decision module to make plausible decisions on them by LSTM, respectively. Fig. 2 shows the architecture of our RDN in details.

Shape-Reasoning Module: To exploit the nonlinear relationship among pairs of states and actions, we leverage a deep architecture equipped with a series of hierarchical nonlinear functions. In practice, the deep network is fed with the observed raw patches \mathbf{s} and then outputs an action $\hat{\mathbf{a}}$ over the continuous space:

$$\hat{\mathbf{a}} = \text{CNN}_{\theta_\pi}(\mathbf{s}) = \text{CNN}_{\theta_\pi}(o(\mathbf{I}, \mathbf{p})), \quad (3)$$

where $\text{CNN}_{\theta_\pi}(\cdot)$ is parameterized by layers of weights and biases. Taking advantages of CNN, the policy learning process is optimized directly starting from raw pixels in an end-to-end fashion, which avoids falling into the sub-optima rather than the separately two-stage manner.

Based on (3), our architecture reasons a shape over the continuous shape space: $\hat{\mathbf{p}} = \mathbf{p} + \hat{\mathbf{a}}$. However, directly generating an unseen shape $\hat{\mathbf{p}}$ over the continuous shape space \mathcal{P} may introduce bias and densely generating shape samples during training process consumes much computational time. To address both problems, we create a discrete shape set \mathcal{D} and then reason a plausible facial shape with the set \mathcal{D} . The shapes from the set \mathcal{D} are produced from training samples by Procrustes analysis [18] (see more details in Sec. 4.3), providing prior and complementary information of shape variances caused by identities and expressions [26], [28], [40]. Note that the reasoned shape $\hat{\mathbf{p}}$ is usually not included in the discrete shape space \mathcal{D} . Having obtained the reasoned shape $\hat{\mathbf{p}}$, our architecture aims to produce a subset of shape candidates within the set \mathcal{D} so as to remove the outliers. Since the MDP agent requires to make decisions on actions, we define a K -nearest neighbour searching function $g_K(\cdot)$ to map $\hat{\mathbf{p}}$ to a subset of discrete actions. Given K , the function $g_K(\cdot)$ finds the K nearest shapes to the reasoned $\hat{\mathbf{p}}$ as the shape candidates. According to $\hat{\mathbf{p}}$, it computes the shape residual for each candidate as the refined action \mathbf{a} . As a result, these obtained actions are the output of $g_K(\cdot)$ and will be evaluated by the proposed shape-decision module, where $g_K(\cdot)$ is defined as follows:

$$g_K(\hat{\mathbf{p}}) = \arg \min_{\mathbf{a}=\hat{\mathbf{p}}-\mathbf{p}} \| \hat{\mathbf{p}} - \mathbf{p} \|_2^2, \quad \mathbf{p} \in \mathcal{D}, \quad (4)$$

where $\| \cdot \|_2$ specifies the ℓ_2 norm to measure the point-to-point distance and K denotes the number of neighbouring shape candidates, respectively.

Shape-Decision Module: Cascaded regression-based methods [3], [9], [29], [53], [69], [78], [85] can be seen as a special policy gradient [79], where the policy is derived as a sequence of feature-to-shape mapping functions and optimized by a common

supervised learning manner. However, the available supervision signals in existing methods are basically hand-designed, so that they are difficult to accurately measure the quality value for the real-world shape. To address this, we employ the MDP state-action value function as the alignment error function for shape evaluation, which achieves the expected total discounted shape quality values. Starting from a given state \mathbf{s}^t and taking an action \mathbf{a}^t under the policy π thereafter, we define the shape evaluation function by the following Bellman equation:

$$Q_\pi(\mathbf{s}^t, \mathbf{a}^t) = \mathbb{E}[r(\mathbf{s}^t, \mathbf{a}^t) + \gamma \cdot Q_\pi(\mathbf{s}^{t+1}, f_{\theta_\pi}(\mathbf{s}^{t+1}))]. \quad (5)$$

Instead of simply using a linear value function, we propose a deep neural networks architecture to approximate the value function as $Q_\pi(\mathbf{s}, \mathbf{a}) \approx \mu_{\theta_Q}(\mathbf{s}, \mathbf{a})$, where θ_Q denotes the network parameters. With the designed networks, we evaluate the quality value for each candidate and make the decision under the policy π (we omit the index t for simplicity):

$$\pi_\theta(\mathbf{s}) = \arg \max_{\mathbf{a} \in g_K \circ f_{\theta_\pi}(\mathbf{s})} \mu_{\theta_Q}(\mathbf{s}, \mathbf{a}), \quad (6)$$

where θ denotes the network parameters of the whole RDN architecture. The output of $g_K \circ f_{\theta_\pi}(\mathbf{s})$ is a subset of actions obtained by the searching method $g_K(\hat{\mathbf{a}})$ and the deterministic policy function $\hat{\mathbf{a}} = f_{\theta_\pi}(\mathbf{s})$.

To semantically exploit the action-value function μ_{θ_Q} , we carefully design a LSTM architecture, which aims to evaluate the shape quality by considering the facial parts individually. Specifically, we partition all landmarks into C parts. As a result, the observed patches are segmented as C parts. To clarify each part, we let ϕ_i denote the i th part-based representation which consists of local patches within each facial part. For each ϕ_i , we design C CNNs to encode the part-based patches as the immediate representation: $\mathbf{h}_i = \text{CNN}_{\theta_Q}(\phi_i) \in \mathbb{R}^H$, where CNN_{θ_Q} specifies the CNN architecture. Hence, we obtain a set of hidden features $[\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_C] \in \mathbb{R}^{H \times C}$ for total facial parts. Taking the global shape constraint into account, we sequentially feed these hidden features to the LSTM and compute a real value to measure the shape quality via the following equation:

$$V = \mu_{\theta_Q}(\mathbf{s}, \mathbf{a}) = \mathbf{u}^T \tanh(\text{LSTM}_{\theta_Q}(\mathbf{h}_i)), \quad (7)$$

where $\text{LSTM}_{\theta_Q}(\cdot)$ denotes the LSTM architecture, and $\mathbf{u} \in \mathbb{R}^H$ is leveraged to map the hidden feature to the shape quality value, and $\tanh(\cdot)$ denotes the nonlinear activation function, respectively.

Inference of Face Alignment: During the testing phase, our RDN model starts with each given facial image \mathbf{I} with an arbitrary initial shape \mathbf{p}^0 . Our architecture produces a shape residual $\hat{\mathbf{a}}$ and then maps it to a subset of shape candidates. Having obtain these candidates, we evaluate them and choose the highest-scoring one as the initialization for future iteration. The whole reasoning-decision execution terminates if the decision values reach the acceptance of the experienced point-to-point distance. With this, we use the reasoned shape $\hat{\mathbf{p}}$ at the final iteration as the resulting outcome. Fig. 2 takes the t -th iteration as an example during the inference process. Note that the iteration number in our RDN is determined by the degree of difficulty cases for testing faces, e.g., less for near-frontal samples and more for profile samples, while the cascaded regression uses a constant iteration number for each given sample. In our experiments, we figure out that our RDN proceeds averaging four iterations for all testing samples and each iteration requires around 10 ms on CPU (refer to more runtime

Algorithm 1: RDN

```

Input: Training set:  $\{(\mathbf{I}_i, \mathbf{p}_i^*)\}_{i=1}^N$  with  $N$  samples,
         $\Lambda = 5, \Gamma = 10$ .
Output: Network parameters:  $\theta_\pi$  and  $\theta_Q$ .
1 Randomly initialize networks  $f_{\theta_\pi}$  and  $\mu_{\theta_Q}$ ;
2 Initialize target networks  $f_{\theta_\pi}$  and  $\mu_{\theta_Q}$  with weights
     $\theta'_\pi \leftarrow \theta_\pi$  and  $\theta'_Q \leftarrow \theta_Q$ ; Initialize replay buffer  $\mathcal{B}$ .
3 for Epoch = 1, ...,  $\Lambda$  do
4   for each training sample  $i = 1, \dots, N$  do
5     Receive initial observation state  $\mathbf{s}^1$ ;
6     for  $t = 1, \dots, \Gamma$  do
7       Select action  $\mathbf{a}^t = \pi_\theta(\mathbf{s}^t)$  by (6);
8       Execute action  $\mathbf{a}^t$  to observe reward  $r^t$  and
         new state  $\mathbf{s}^{t+1}$ ;
9       Store transition  $(\mathbf{s}^t, \mathbf{a}^t, r^t, \mathbf{s}^{t+1})$  in  $\mathcal{B}$ ;
10      Randomly sample a mini-batch of transitions
           $(\mathbf{s}^t, \mathbf{a}^t, r^t, \mathbf{s}^{t+1})$  from  $\mathcal{B}$ ;
11      Set  $\mathbf{a}^{t+1} = \pi_\theta(\mathbf{s}^{t+1})$  and  $y^t$  via (10);
12      Optimize  $f_{\theta_Q}$  by minimizing the loss function
          (9);
13      Update parameters of  $f_{\theta_\pi}$  by (8) and the target
          network parameters.
14   end
15 end
16 end
17 Return: Network parameters  $\theta_\pi$  and  $\theta_Q$ .

```

details in Section 5), which exploits the promising performance in terms of both the accuracy and efficiency.

3.3 Optimization via Policy Gradient

The whole optimization process executes by following [38], [62], which typically adjusts the policy gradient by maximizing shape quality values. Specifically, to optimize (2), we collect all weights of the CNN_{θ_π} for the parameters of the shape-reasoning f_{θ_π} . Based on the deterministic policy gradient theorem [62], we compute the gradient with respect to (2) as follows:

$$\begin{aligned} \nabla_{\theta_\pi} J &= \int_{\mathcal{S}} \rho_\pi(\mathbf{s}) \nabla_{\theta_\pi} \mu_{\theta_Q}(\mathbf{s}, \mathbf{a})|_{\mathbf{a}=f_{\theta_\pi}(\mathbf{s})} d\mathbf{s} \\ &= \mathbb{E}_{\mathbf{s} \sim \rho_\pi} \left[\nabla_{\mathbf{a}} \mu_{\theta_Q}(\mathbf{s}, \mathbf{a})|_{\mathbf{a}=f_{\theta_\pi}(\mathbf{s})} \cdot \nabla_{\theta_\pi} f_{\theta_\pi}(\mathbf{s})|_{\mathbf{s}} \right] \\ &\approx \frac{1}{N} \sum_i \nabla_{\mathbf{a}} \mu_{\theta_Q}(\mathbf{s}, \mathbf{a})|_{\mathbf{a}=f_{\theta_\pi}(\mathbf{s}^i)} \cdot \nabla_{\theta_\pi} f_{\theta_\pi}(\mathbf{s})|_{\mathbf{s}^i}. \end{aligned} \quad (8)$$

Although the K -NN searching component in our RDN is not fully differentiable, we consider that the function g_K as a part of the deterministic policy function f_{θ_π} . This allows us to train f_{θ_π} on its own output $\mathbf{a} = \hat{\mathbf{a}}$ in (8) by following [15], [38].

Accordingly, to optimize (6) in our shape-decision module, we collect all parameters of $\text{CNN}_{\theta_Q}(\cdot)$, $\text{LSTM}_{\theta_Q}(\cdot)$ and \mathbf{u} in (7) as the parameters of μ_{θ_Q} . Specifically, these parameters are optimized via Q-learning [71], [72] and the loss function is formulated as minimizing the following optimization problem as follows:

$$J(\theta_Q) = \mathbb{E}_{\mathbf{s}, \mathbf{a}, t} [(\mu_{\theta_Q}(\mathbf{s}^t, \mathbf{a}^t) - y^t)^2], \quad (9)$$

where the target value is computed based on (5) as

$$y^t = r(\mathbf{s}^t, \mathbf{a}^t) + \gamma \cdot \mu_{\theta_Q}(\mathbf{s}^{t+1}, \mathbf{a}^{t+1}). \quad (10)$$

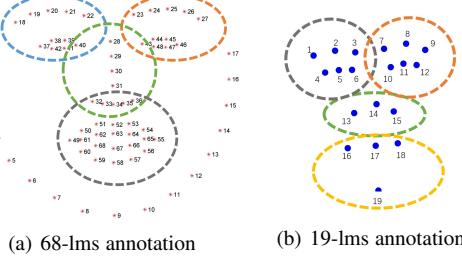


Fig. 3. Landmark partition of two specific types, *i.e.*, the 300-W and AFLW-Full datasets for standard annotations. Note that we did not plot the facial contour for better visualization in the 68-lms annotation.

Since the expectation in (2) depends only on the state space, it is possible to use an off-policy learning method [62] to obtain the function f_{θ_π} . In the context of [38], we also employ a replay strategy and separated target networks for calculating the target value y^t in (10) softly. Moreover, we initialize the target networks by creating a copy of the networks of θ_π and θ_Q , and the parameters are learned by the update equation: $\theta'_{\pi,Q} \leftarrow \tau\theta_{\pi,Q} + (1-\tau)\theta'_{\pi,Q}$. Besides, for an efficient replay strategy, we employ a framework for prioritizing experience [58], which aims to replay significant transitions more frequently for fast convergence.

Algorithm 1 details the training procedure of our RDN.

4 IMPLEMENTATION DETAILS

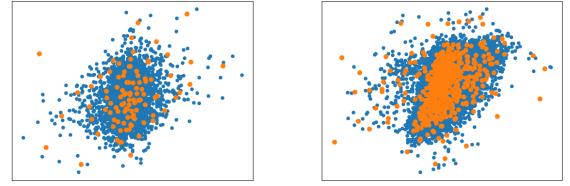
In this section, we describe the implementation details with respect to our proposed RDN including the specifications of proposed network architecture, landmark partition strategy for shape evaluation and the Set \mathcal{D} construction process.

4.1 Network Specifications

Our model was built based on the popular accelerated deep learning toolbox TensorFlow [1], which mainly operates on data flow graphs. For the input data preparation, we first averaged all face boundingboxes across the whole dataset as the output size. Then we rescaled both the detected facial images with padding zeros and the corresponding annotations with the restricted output scales. For each evaluation dataset, the image resolution of the input is determined by the averaged size over all detected faces.

In terms of CNN_{θ_π} in our shape-reasoning module, the first convolution layer accepts L raw local patches in size of 26×26 . Then we deployed two convolutional layers (3×3 kernel size, 1×1 stride) with 64 and 128 kernels. By following the convolution layers, we appended a two-layer fully connections, where the parameters are 128×256 and $256 \times 2L$ matrices (L is the landmark number, *i.e.*, 68 landmarks for both 300-W and 300-VW, or 19 landmarks for AFLW-Full). In terms of the shape-decision module, the CNN_{θ_Q} cell is equipped with the similar specification of CNN_{θ_π} , which incorporates two layers of 32 and 64 kernels with 3×3 convolution, 1×1 stride, one fully connected layer parameterized by 64×128 ($H = 128$). We also employed ReLU as the nonlinear activation function in both CNNs.

K in the K -NN component was chosen as 4 by a cross validation procedure. ϵ in (1) was empirically specified to 0.02. For the hyper-parameters during the training process, we specified the learning rates for both target networks to 0.001, the discounted factor to 0.9 and the smoothness τ to 0.01, respectively. Besides, we sampled 120 transitions in the replay buffer \mathcal{B} . It is valuable



(a) 100 centers in 300-W (b) 1000 centers in AFLW-Full

Fig. 4. Visualization of the clustering centers for both evaluation datasets (300-W and AFLW-Full), where orange points specify clusters and blue ones for all training samples pre-processed by Procrustes [18] analysis.

to notified that we first learned a simple policy f_{θ_π} for our shape-reasoning module CNN_{θ_π} by using the training samples, and then tuned both network parameters θ_π and θ_Q in an end-to-end manner. The whole training procedure converged until the validation error remained minimized and unchanged. Besides, we randomly oversampled entire face images during training process by horizontal flipping and randomly rotation with $\pm 15^\circ$ to augment the training data [68].

4.2 Landmark Partition for Shape Evaluation

In this section, we presented the implementation details of the landmark partition in our proposed shape-decision module. Inspired by [5], [85], we divided the whole face into several parts and then evaluated each part individually. To efficiently evaluate each shape candidate with the facial appearance, we carefully designed a unified CNN-LSTM architecture, which aims to criticize individual facial parts with CNN and moreover takes the global facial shape constraint into account by using LSTM. Due to two types of the widely used facial landmark annotations (*i.e.*, 68 landmarks for 300-W/300-VW, respectively, 19 landmarks for AFLW-Full), we typically operated the landmark partition in two separated manners, accordingly. To be specific, for the dense 68-annotation setting, we partitioned each shape candidate by $C = 5$ overlapped facial parts, facial contour (1-17), left eye (18-22, 37-40), right eye (23-27, 43-46), nose (28-36, 40, 43), mouth (32-36, 49-68). More details are illustrated in Fig. 3(a). Note that we did not plot the facial contour for better visualization under 68-lms annotations. For the sparse 19-landmark annotation setting, we partitioned each shape candidate by *non-overlapping* facial parts $C = 4$ for AFLW-Full, *i.e.*, left eye (1-6), right eye (7-12), nose (13-15), mouth (16-19). Besides, Fig. 3(b) shows the 19-landmark partition employed in our experiments.

4.3 Construction of the Set \mathcal{D}

Before creating the Set \mathcal{D} , we used the Procrustes analysis [18] on all shapes in the training set, which typically normalizes these shapes with respect to a global similarity transformation. By doing this, the effects of scaling, rotation and translation are removed for better shape reasoning. Having obtained these normalized shapes, we applied the clustering algorithm on them and obtained the clustered centers, *e.g.*, 100 cluster centers for 300-W dataset. Finally, we directly regarded these clustered centers as the elements in the constructed Set \mathcal{D} , which provides the shape basis for shape generation, as visualized in Fig. 4. For the AFLW-Full dataset, we conducted 1000 centers based on the existing training samples. This is because samples in the AFLW-Full dataset undergo intense

pose variances. For the annotation in 300-VW which is similar with 300-W, we directly used the Set for 300-VW evaluation. It is valuable to notified that we also normalized the facial images by this global transformation via Procrustes analysis [18] with respect to the normalized ground-truth shapes.

5 EXPERIMENTS

In this section, we present the used dataset description, evaluation protocols and settings, experimental results and model analysis.

5.1 Datasets

For experimental evaluation, we evaluated our proposed RDN approach on the widely used 300-W [55], 300-VW Category 3 [59] and AFLW-Full [30], [86] face alignment datasets compared with existing state-of-the-art methods.

300-W [55]: The 300-W dataset mainly contains three subsets including LFPW [6], HELEN [35] and IBUG [56], where these face samples were captured in unconstrained environments. Following the standard settings in [55], we utilized the training sets of LFPW and HELEN to learn our model. Then we evaluated our RDN on the 224-image LFPW testing set, the 330-image HELEN testing set, as well as the 135-image IBUG, accordingly. We also investigated our approach by following another wildly used evaluation setting: using the LFPW and HELEN testing set as Commonset (554-image), the 135-image IBUG dataset as Challengingset (135-image), and the union of them as Fullset (689-image), respectively. For all experiments in 300-W dataset, our model was trained on the 3148 images of the training set with 68-point markup and the bounding boxes produced by the in-house face detector from the 300-W website.

COFW [8]: The Caltech Occluded Face in the Wild dataset, dubbed COFW, collected from the Internet, consists of 1345 training face images and the remaining 507 face images are utilized for testing. It is notified that most of faces undergo severe occlusion issue due to variations of facial poses, wearing glasses and partial object beyond faces. All face images were manually annotated with 29 landmarks together with the auxiliary visibility/invisibility information (not used in this work) to exploit the occlusion.

300-VW Category 3 [59]: The 300-VW Category 3 dataset is derived from the 300-VW [59] dataset, which was collected specifically for video-based face alignment. The Category 3 in 300-VW contains 14 videos in severe wild conditions and each video lasts around one minute (25-30 images per second). To train our RDN, we employed both the 300-W [55] training images and thousands of samples uniformly drawn from the 300-VW training set. According to the experimental protocol from the 300-VW organizer [59], we used the indices of frames that removed for evaluation.

AFLW-Full [30], [86]: The AFLW-Full dataset contains 24386 images and most of them undergo a large variety in appearance including intense large poses, *e.g.*, near-frontal and profile ($> 60^\circ$) face samples. By following the setting in [86] with only 19 re-annotated landmarks, we used 20, 000 images for training and the remaining 4, 386 images for testing for evaluation.

5.2 Standard Evaluation Protocols

We employed the two types of comparisons with averaged error and cumulative error for evaluation.

Averaged Error Comparisons: We leveraged the standard normalized root mean squared error (NRMSE) [55], which computes the mean error normalized by the inter-pupil distance as follows:

$$\text{NRMSE} = \frac{\sum_{i=1}^L \|\hat{p}_i - p_i^*\|_2}{L \cdot \|p_{l\text{eye}}^* - p_{r\text{eye}}^*\|_2} \quad (11)$$

where $\|\cdot\|_2$ denotes the ℓ_2 norm, \hat{p}_i and p_i^* denote the i -th landmark coordinates of the predicted and ground-truth facial landmark positions accordingly. $p_{l\text{eye}}$ and $p_{r\text{eye}}$ denote the pupil locations (eye centers) of the left eye and the right eye, respectively. For the 300-W dataset, we employed the inter-pupil distance as the normalization term, which is different from the standard 300-W protocol that uses outer eye corner distance. To be specific, we followed the evaluation protocol widely employed in ESR [9] and CFSS [85] to use the "inter-pupil" (eye-center-distance) distance as the normalization factor. Having obtained the NRMSE for each testing image, we averaged the NRMSE errors over total testing images for the averaged error comparisons. Due to various facial profiles on the AFLW-Full dataset, we normalized the point-to-point distance by the square root of annotated boundingbox size [26], [86].

CED Curves: We also conducted the cumulative error distribution (CED) curves [55] of NRMSE errors to quantitatively evaluate the performance in comparisons to the state-of-the-arts. The CED is computed by the following equation:

$$\text{CED} = \frac{N_{e \leq l}}{N} \quad (12)$$

where $N_{e \leq l}$ is the number of images on which the error l is no less than e .

For averaged error comparisons, the performance is higher when the errors are lower, and the performance is lower, vice versa. Contrastively, for the CED curve comparisons, the performance is higher when the curves accordingly are higher. Both evaluation protocols are widely used in the standard face alignment benchmarking datasets [30], [55], [59].

5.3 Comparison with State-of-the-arts

We compared our RDN with conventional hand-crafted methods [3], [8], [29], [53], [69], [78], [85], [87] and deep learning-based methods [41], [44], [57], [68], [83], [84]. For fair comparisons, we re-produced the results with their publicly released codes. For those without source codes, we directly cropped the reported results from their original papers.

Evaluation on 300-W: Fig. 5(a), 5(b), 5(c) show the comparisons of CED curves on the 300-W LFPW, HELEN and Challengingset (IBUG) with the standard evaluation protocol, respectively. We see that our RDN consistently obtains higher performance than the state-of-the-arts. Moreover, our RDN even outperforms the deep learning-based methods [41], [44], [57], [68], [83], [84] by a large margin in the 300-W Challengingset. This is because our approach learns to evaluate the predicted shapes and re-initialized the following-iteration initial shape via deep reinforcement learning. Hence, an optimal shape searching path is globally mined over the continuous shape space, which maximizes the shape quality values. In Table 1, we also reported the averaged error comparisons on 300-W Commonset, Challengingset and Fullset under another evaluation setting. We see that our RDN obtains lower averaged errors compared with the variants of cascaded

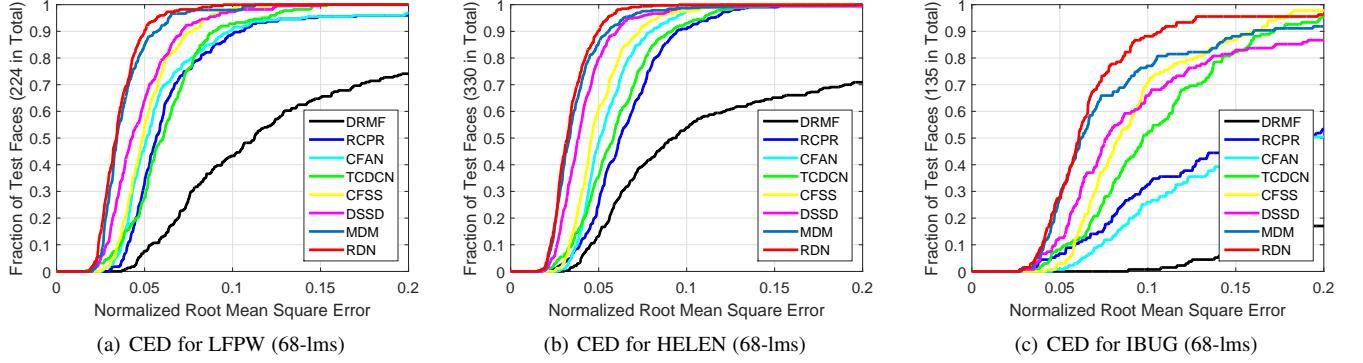


Fig. 5. CED curves on 300-W including LFW, HELEN, IBUG, respectively, where 68 landmarks were employed for evaluation. Our RDN model outperforms state-of-the-arts consistently on three datasets, which indicates the robustness of our RDN to large variations due to types of facial poses, diverse facial expressions, partial occlusions in wild conditions. Note that the comparison curves should be best viewed in the color pdf file.

TABLE 1

Comparisons of averaged errors (100%) normalized by eye-center-distance of our RDN with existing methods (in chronological order) on the 300-W Commonset, Challengingset and Fullset, respectively, where 68 landmarks were employed for evaluation.

Method	Commonset	Challengingset	Fullset
Zhu <i>et.al</i>	8.29	8.16	18.33
FPLL [87]	8.22	18.33	10.20
DRMF [3]	6.65	19.79	9.22
RCPR [8]	6.18	17.26	8.35
SDM [78]	5.57	15.40	7.50
ESR [9]	5.28	17.00	7.58
LBF [53]	4.95	11.98	6.32
Deep Reg [60]	4.51	13.80	6.31
CFAN [83]	5.50	16.78	7.69
CFSS [85]	4.73	9.98	5.76
TCDCN [84]	4.80	8.60	5.54
RAR [77]	4.12	8.35	4.94
MDM [68]	4.83	10.14	5.88
DSSD [41]	4.16	9.20	5.59
DeepReg [44]	4.36	7.42	4.96
TCD [33]	3.67	7.62	4.44
CPR [14]	3.39	8.14	4.36
RDN	3.31	7.04	4.23

shape regression-based. Since our RDN applies reinforcement learning to address face alignment as the MDP, the optimal MDP policy is achieved by maximizing the cumulated reward from the MDP. Hence, high performance has been obtained due to the benefit of making decisions on finding a plausible shape search path over the continuous shape space.

Fig. 8 shows the comparisons of our approach with the general cascaded regression with deep neural networks (we use MDM [68] composed by four stages as the baseline). Each sample specifies the plotting immediate results for different iterations. For our RDN, we visualized the results of the shape-reasoning module \hat{p} and those of the shape-decision module p . The termination represents the final results for \hat{p} of the last iteration. Based on the results, we made two-fold conclusions: 1) The cascaded regression-based methods such as MDM, usually lead to misalignment errors even at the earlier cascaded executions, which results in bias prediction (sub-optima) in the latter stages cumulatively. 2)

TABLE 2

Comparisons of AUC and Failure Rate of our approach with recent compelling methods on 300-W Testset (68-lms). Note the accuracy is reported as the AUC and Failure Rate (*calculated at a threshold of 0.1*).

Methods (%)	AUC	Failure Rate
Deng <i>et al.</i> [12]	47.52	5.5
Fan <i>et al.</i> [16]	48.02	14.83
MDM [68]	45.32	4.03
DenseReg + MDM [19]	52.19	3.67
JMFA [13]	54.85	1.00
LAB [74] [†]	58.85	0.83
RDN	53.7	2.43

[†] At the time writing, we have not directly made access to the results in LAB [74] and moreover taken full advantages of LAB [74] in our experimental comparisons.

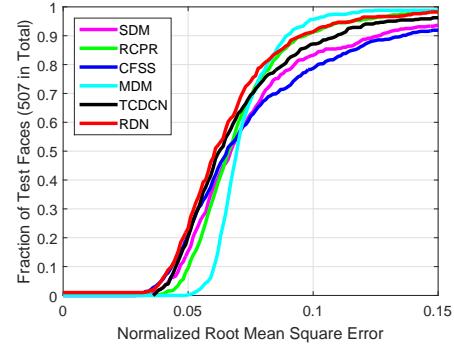


Fig. 6. CED curves on the COFW dataset principally regarding of the occlusion issue, where 29 landmarks were employed for evaluation. We fix the error-axis within the range of [0,0.15] in focus on significant comparisons. Note that the comparison curves should be best viewed in the color pdf file.

However, our RDN can reason a sequence of plausible shapes with the Set \mathcal{D} and re-initialize the shape for future iteration under such cases when an arbitrary initial shape was given. This is because we propose a shape-reasoning policy by evaluating the shape values under the deep reinforcement learning framework. Thus, our RDN exploits to mine a plausible shape searching path globally over the shape space which maximizes the shape quality values across iterations.

TABLE 3

Comparisons of the averaged errors with other approaches on COFW regarding of occlusions, where 29 landmarks were employed for evaluation.

Method	FPLL [87]	RCPR [8]	HPM [17]	RPP [80]	SDM [78]	CFAN [83]	TCDCN [84]	MDM [68]	PCD [33]	RDN
Error	8.79	8.50	7.46	7.52	8.77	8.38	8.05	6.26	6.02	5.82

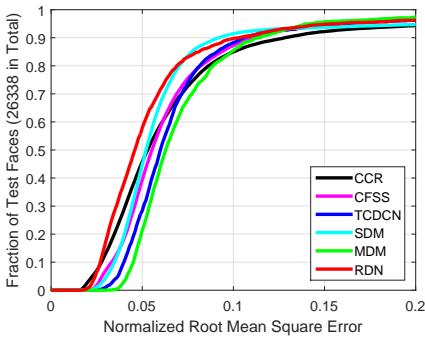


Fig. 7. CED curves on the 300-VW challenging set Category 3, which demonstrates the robustness of our RDN versus temporal face changes. We employed 68 landmarks for evaluation as the same setting in 300-W. Note that the comparison curves are best viewed in the color pdf file.

As discussed in MDM [68], there is no consistent way of the normalized factors for alignment, *e.g.*, inter-ocular distance (300-W) or inter-pupil distance. To clarify the results, we compute the area-under-the-curve (AUC) and failure rate (error greater than 0.1 as a failure) on the 300-W Testset. The 300-W Testset is the private test-set used for the 300-W competition and recently made publicly available, which is quite different from the 300-W Fullset. Specifically, this dataset consists of 600 images and is divided into the indoor and outdoor subsets. As shown in Table 2, our RDN achieves very competitive performance compared with more recent face alignment methods such as JMFA [13] and DenseReg + MDM [19]. However, LAB [74] was equipped with a deeper hourglass-based [46] architecture, which contains more parameters than ours and may be trapped into degrading efficiency performance especially with limited computing resources, *e.g.*, on a single CPU.

Evaluation on COFW: To investigate the effectiveness of our methods versus various occlusions on the COFW dataset, where the facial images are occluded by the invisible parts. Table 3 shows the averaged errors (Note that we skip the metric of failure rates [8] for comparisons, because our learned objective function mainly focuses on averaged mean error instead of detection accuracy of occlusions, which can be the future work based on our architecture). From these results, we see that our method achieves better performance than that of the state-of-the-art methods. Some sample alignment results on the COFW dataset are shown in Fig. 9. Obviously, our model achieves the robustness to the face images versus partial occlusions by exploiting both the local and global information in the learned shape evaluation function, which is also back-propagated in the reasoning descents via reinforcement learning. The main reason lies that we adjust the gradient in the directions of the globally-optimized shape evaluation function.

Evaluation on AFLW-Full: To further evaluate the superiority of our approach, we conducted experimental comparisons on the AFLW-Full [86] dataset regarding of severe large-pose

cases such as near-frontal and profile faces. Table 4 tabulates the averaged error comparisons with state-of-the-art methods [8], [29], [44], [53], [69], [78], [85], [86]. According to these results, we observe that our approach outperforms the cascaded regression-based methods like SDM, ERT and MDM, and moreover performs higher performance than the multi-view approaches such as CFSS and CCL. Moreover, Fig. 11 shows some selected challenging face samples under our framework. As can be inferred, our approach achieves robustness to intense facial poses compared with existing approaches. This is because our RDN learns to refine the shape searching path by evaluating the shape quality with the proposed deep reinforcement learning framework. As a result, for each face with an extreme pose, our architecture selects plausible actions to remove the outlier shape candidates, which efficiently prevents our model from being trapped into the sub-optima.

Evaluation on 300-VW Category 3: We evaluated our proposed RDN on the challenging 300-VW Category 3 dataset [59] compared with both the image-based and video-based face alignment methods. For the image-based methods [78], [84], [85], we conducted experiments with their publicly source codes in a tracking-by-detection protocol. For CCR [57] that was designed for video-based face alignment, we directly used the released results from the authors for fair comparisons. It is worth noting that our method obtains comparable performance in comparisons to [48], [57], [59] and we did not include the results due to its saturated performance [20], [42]. Fig. 7 and Fig. 10 show the quantitative CED curves of our approach compared with most of state-of-the-art methods and qualitative results, respectively. The results indicate that our approach achieves better results than the state-of-the-arts, and even outperforms the video-based method CCR [57] which exploits the temporal information in their model. Since our image-based RDN architecture is complementary to the video-based face alignment approaches, we believe that our RDN architecture can receive a significant improvement on the alignment performance by combining these temporal modeling techniques [42], [48], [57].

5.4 Analysis

In this subsection, we performed an ablation study for different components of our RDN architecture.

Investigation of Different Components in RDN: We studied the importance of the designed components in our RDN architecture. Specifically, we defined the following baseline methods: 1) *w/o shape quality evaluation* to directly use the equation (1) as the target value function, and 2) *w/o K-NN* to directly evaluate the reasoned action \hat{a} by the equation (7). Fig. 13 demonstrates the comparisons of the normalized reward values with both baselines. From the results, the baseline 1) *w/o shape quality evaluation* degrades the performance, which is because the shape quality evaluation principally adjusts the policy gradient to find an optimal shape searching path. Moreover, we obtain better results than the baseline 2) *w/o K-NN*. The reason lies that the prepared shape clusters exploiting prior information to exploit efficiency on shape decision process. In summary, our RDN exploits the effectiveness

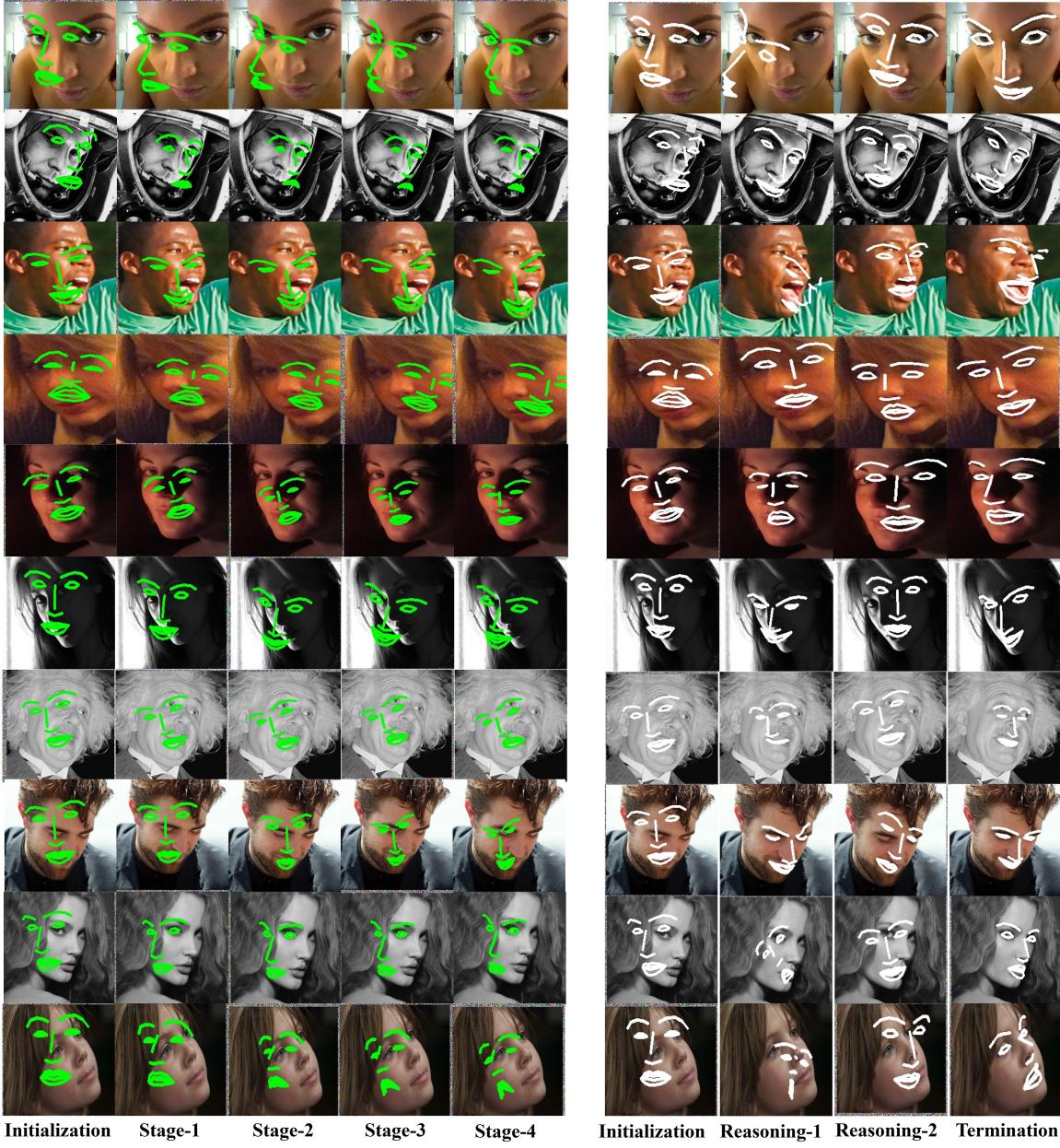


Fig. 8. Representative results of the comparisons of our RDN approach (white lines) versus the conventional cascaded regression approach (green lines). We used MDM [68] as the baseline cascaded regression approach with deep neural networks and plotted the immediate outcomes for each stage. For our RDN, we plotted the outcomes of both the shape-reasoning and shape-decision networks. Note that even starting with a random and low-quality initialization drawn from the training samples, our RDN achieves more reasonable initialization after plausible reasoning and evaluation.

integrated with both the shape evaluation strategy and the discrete shape prior information.

Investigation of the Robustness to Initializations: To investigate the robustness of our RDN regarding with arbitrary initialization, we compared two types of state-of-the-art approaches including CFSS [85] with multi-view initializations and MDM [68] with mean-shape initialization. Specifically, we first removed the groundtruth shape of the testing image from the training set and then randomly sampled one from the remaining shapes as its initialization. Table 5 tabulates the comparisons of averaged errors on 300-W Fullset. Compared with multi-view initialization, *i.e.*, CFSS, we obtain lower NRMSE errors. Moreover, the performance of our RDN degrades slightly with arbitrary initialization. In addition, Fig. 12 demonstrates some selected challenging samples

from 300-W Challengingset. We see that our model promisingly achieves robustness to diverse facial variances such as severe poses and exaggerated expressions. As can be inferred from these results that our approach acts for reasonable inference and then plausibly selects robust initialization. Even with an low-quality initialization, our proposed model significantly outperforms other approaches on a volume of challenging cases including extreme poses, exaggerate expressions and severe occlusions. Therefore, our architecture achieves an optimal shape searching path over the global shape space even starting from an arbitrary initialization.

Investigation of Estimated Shape Quality Values: Estimating the shape quality value V is a key step in our RDN. To evaluate the effectiveness of the expected shape quality values, we compared the expected versus estimated values V which were

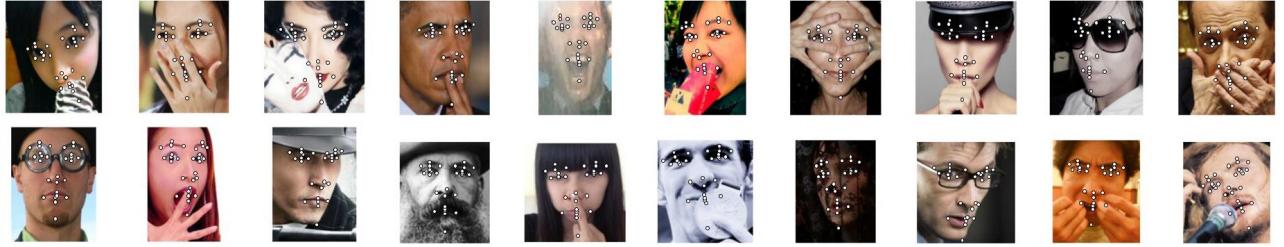


Fig. 9. Representative results of our proposed RDN approach on the COFW dataset, where 29 landmarks were employed for evaluation. According to the qualitative analysis, we see that our RDN even achieves robust performance regarding with diverse partial occlusions.



Fig. 10. Representative results of our proposed RDN approach on the 300-VW dataset, where 68 landmarks were employed for evaluation. According to the qualitative analysis, we see that our RDN still achieves robust performance regarding with temporal changes due to head motions.

TABLE 4

Comparisons of averaged errors on ALFW-Full (in chronological order, 19-lms). We achieve very competitive performance versus other methods.

Method	RCPR [8]	SDM [78]	ERT [29]	PO-CR [69]	CFSS [85]	CCL [86]	MDM [68]	DeepReg [44]	CPR [14]	RDN
Error	3.73	4.05	4.35	5.32	3.92	2.72	2.53	2.12	2.33	2.06

TABLE 5

Performance effects of our approach with the mean shape versus randomly initialized shapes on the 300-W Fullset. We consider CFSS and MDM as the baselines. Note that we repeated the procedure for five times and then averaged total results (marked with '**'). We see that our model achieves robust performance versus arbitrary initializations.

Method	Initialization Type	Averaged Error
CFSS [85]	Multiple-View	5.76
MDM [68]	Arbitrary Shape	8.75**
MDM [68]	Mean Shape	5.59
Ours	Arbitrary Shape	4.52
Ours	Mean Shape	4.23**

sampled across iterations during the testing phase. As shown in Fig. 14, the plotting points stand for the ratios between the estimated and expected values whereas the dashed line denotes the ground-truth (they should be ideally equivalent), respectively. The results shows the effectiveness of the learned function μ_{θ_Q} which aims to evaluate the shape quality values. Moreover, it can be seen that our model converges motivated by the fact of all points approaching the dashed line.

Time Complexity Analysis: Our RDN is composed by the shape reasoning module and the shape decision module (Section 3). The computational complexities for both modules basically depend on the designed deep neural networks architectures

which are denoted by f_{θ_π} and μ_{θ_Q} . Let $\mathcal{O}(f_{\theta_\pi})$ be the complexity of the shape-reasoning module, K denote the neighbour searching number for shape candidates (we empirically assigned K to 4 in our experiments), $\mathcal{O}(\mu_{\theta_Q})$ denote that of the shape-decision module. We suppose the required iteration number for each testing image is denoted by \mathcal{I} . The time complexity of each inference sample is derived as $\mathcal{I} \cdot (\mathcal{O}(f_{\theta_\pi}) + K \cdot \mathcal{O}(\mu_{\theta_Q}))$. To better exploit the efficiency, our RDN model can evaluate K shape candidates in a parallel manner, which reduces the time complexity to $\mathcal{I} \cdot (\mathcal{O}(f_{\theta_\pi}) + \mathcal{O}(\mu_{\theta_Q}))$.

6 CONCLUSION AND FUTURE WORK

In this paper, we have proposed a deep reinforcement learning architecture for robust face alignment, which reasons a sequence of plausible descent directions by maximizing the received values across multiple iterations. In parallel, our model learns a globally-optimized function as the alignment error function to evaluate the shape quality. Extensive experimental results on four unconstrained benchmarks show the effectiveness of our proposed approach versus variant poses, expressions, partial occlusions and motion changes in video.

For future works, we are interested in exploiting the frame-context information specific for video-based face alignment [20], [42], [57] by partially-labelled frames for model extension. Moreover, it is desirable to infer a set of collaborate and competitive

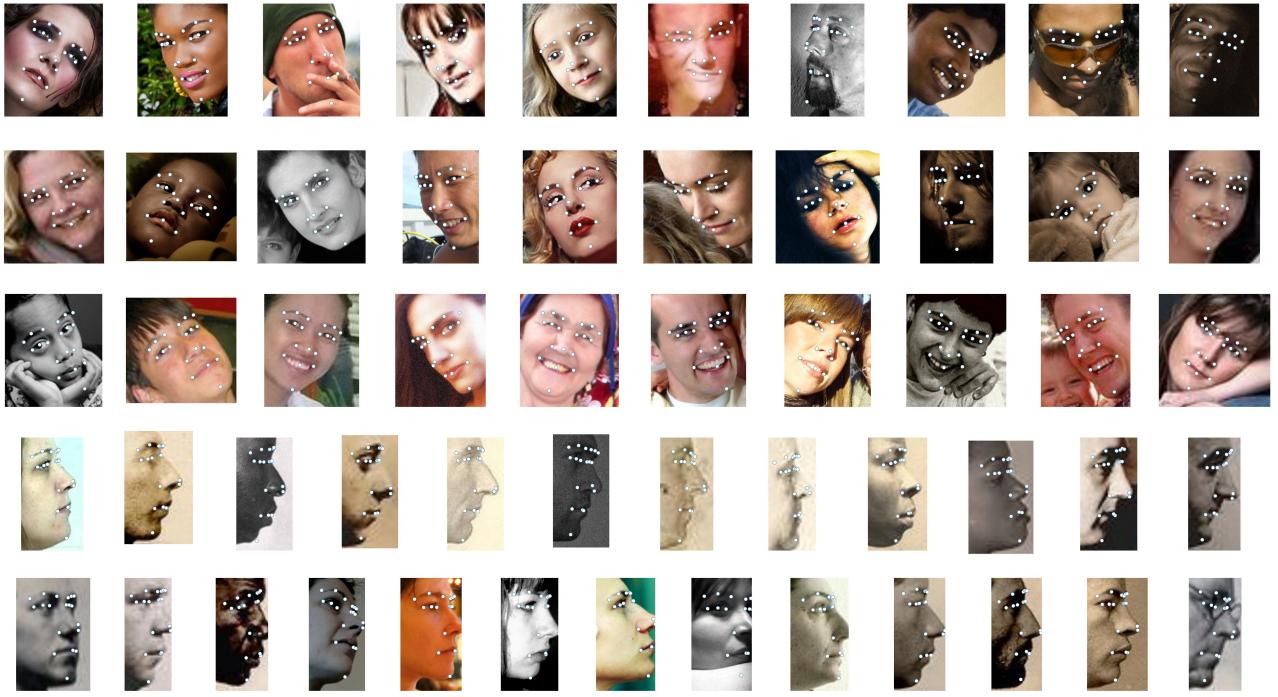


Fig. 11. Representative results of our proposed RDN approach on AFLW-Full versus various poses, where 19 landmarks were employed for evaluation. According to these results, we see that our RDN relatively achieves robust performance versus both the near-frontal and profile faces.



Fig. 12. Representative plotting examples on 300-W Challengingset (68-lms) compared with baselines. First row: general cascaded regression [68] with the mean shape; Second row: coarse-to-fine shape searching [85] with multi-view initializations; Last row: our proposed RDN with arbitrary initializations. Note our approach achieves robust performance versus diverse challenging cases, which exhibits independent of poor initializations.

gradient directions with multi-domain partition [79], [86] under the multi-agent deep reinforcement learning framework [47]. In addition to efficient face alignment, one possible solution is to restrict the testing time tolerance and to plug some runtime deep compression strategies [39] in the meantime of the shape update procedure under the proposed framework.

ACKNOWLEDGEMENTS

This work was supported in part by the National Natural Science Foundation of China under Grant 61806104, Grant 61672306, Grant U1713214, Grant 61572271, Grant 61662059 and Grant 61527808, in part by the National Key Research and Development Program of China under Grant 2017YFA0700802, in part by the National 1000 Young Talents Plan Program, and in part by the Shenzhen Fundamental Research Fund (Subject Arrangement) under Grant JCYJ20170412170602564.

REFERENCES

- [1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [2] K. T. Abou-Moustafa, F. D. la Torre, and F. P. Ferrie. Pareto discriminant analysis. In *CVPR*, pages 3602–3609, 2010.
- [3] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic. Robust discriminative response map fitting with constrained local models. In *CVPR*, pages 3444–3451, 2013.
- [4] T. Baltrusaitis, P. Robinson, and L. Morency. Openface: An open source facial behavior analysis toolkit. In *WACV*, pages 1–10, 2016.
- [5] P. N. Belhumeur, D. W. Jacobs, D. J. Kriegman, and N. Kumar. Localizing parts of faces using a consensus of exemplars. In *CVPR*, pages 545–552, 2011.

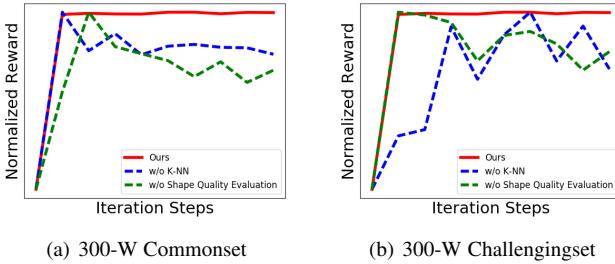


Fig. 13. Normalized rewards of our RDN compared with other baseline methods. These results demonstrate the importance of the proposed components of the shape evaluation function and the K-NN shape searching constraint.

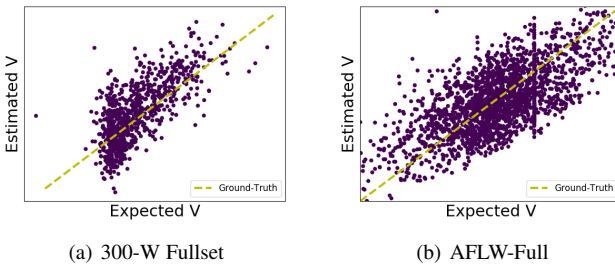


Fig. 14. Evaluation on estimated value versus expected value (refer to V in Section 3). Note that the expected values collect the sampled cumulative reward across training iterations.

- [6] P. N. Belhumeur, D. W. Jacobs, D. J. Kriegman, and N. Kumar. Localizing parts of faces using a consensus of exemplars. In *CVPR*, pages 545–552, 2011.
- [7] A. Bulat and G. Tzimiropoulos. How far are we from solving the 2d and 3d face alignment problem (and a dataset of 230,000 3d facial landmarks). In *ICCV*, Oct 2017.
- [8] X. P. Burgos-Artizzu, P. Perona, and P. Dollár. Robust face landmark estimation under occlusion. In *ICCV*, pages 1513–1520, 2013.
- [9] X. Cao, Y. Wei, F. Wen, and J. Sun. Face alignment by explicit shape regression. In *CVPR*, pages 2887–2894, 2012.
- [10] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *TPAMI*, 23(6):681–685, 2001.
- [11] T. F. Cootes, M. C. Ionita, C. Lindner, and P. Sauer. Robust and accurate shape model fitting using random forest regression voting. In *ECCV*, pages 278–291, 2012.
- [12] J. Deng, Q. Liu, J. Yang, and D. Tao. M^3 CSR: multi-view, multi-scale and multi-component cascade shape regression. *Image Vision Computing*, 47:19–26, 2016.
- [13] J. Deng, G. Trigeorgis, Y. Zhou, and S. Zafeiriou. Joint multi-view face alignment in the wild. *CoRR*, abs/1708.06023, 2017.
- [14] X. Dong, S.-I. Yu, X. Weng, S.-E. Wei, Y. Yang, and Y. Sheikh. Supervision-by-Registration: An unsupervised approach to improve the precision of facial landmark detectors. In *CVPR*, 2018.
- [15] G. Dulac-Arnold, R. Evans, P. Sunehag, and B. Coppin. Reinforcement learning in large discrete action spaces. *CoRR*, abs/1512.07679, 2015.
- [16] H. Fan and E. Zhou. Approaching human level facial landmark localization by deep learning. *Image Vision Computing*, 47:27–35, 2016.
- [17] G. Ghiasi and C. C. Fowlkes. Occlusion coherence: Localizing occluded faces with a hierarchical deformable part model. In *CVPR*, pages 1899–1906, 2014.
- [18] J. C. Gower. Generalized procrustes analysis. *Psychometrika*, 40(1):33–51, Mar 1975.
- [19] R. A. Güler, G. Trigeorgis, E. Antonakos, P. Snape, S. Zafeiriou, and I. Kokkinos. DenseReg: Fully convolutional dense shape regression in-the-wild. In *CVPR*, pages 2614–2623, 2017.
- [20] M. Haris Khan, J. McDonagh, and G. Tzimiropoulos. Synergy between face alignment and tracking via discriminative global consensus optimization. In *ICCV*, Oct 2017.
- [21] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.
- [22] I. Hosu and T. Rebedea. Playing atari games with deep reinforcement learning and human checkpoint replay. *CoRR*, abs/1607.05077, 2016.
- [23] J. Hu, J. Lu, and Y. Tan. Discriminative deep metric learning for face verification in the wild. In *CVPR*, pages 1875–1882, 2014.
- [24] Z. Huang, X. Zhao, S. Shan, R. Wang, and X. Chen. Coupling alignments with recognition for still-to-video face recognition. In *ICCV*, pages 3296–3303, 2013.
- [25] Z. Jie, X. Liang, J. Feng, X. Jin, W. F. Lu, and S. Yan. Tree-structured reinforcement learning for sequential object localization. In *NIPS*, 2016.
- [26] A. Jourabloo and X. Liu. Pose-invariant 3d face alignment. In *ICCV*, pages 3694–3702, 2015.
- [27] A. Jourabloo and X. Liu. Large-pose face alignment via cnn-based dense 3d model fitting. In *CVPR*, pages 4188–4196, 2016.
- [28] A. Jourabloo, M. Ye, X. Liu, and L. Ren. Pose-invariant face alignment with a single cnn. In *ICCV*, Oct 2017.
- [29] V. Kazemi and J. Sullivan. One millisecond face alignment with an ensemble of regression trees. In *CVPR*, pages 1867–1874, 2014.
- [30] M. Köstinger, P. Wohlhart, P. M. Roth, and H. Bischof. Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization. In *ICCVW*, pages 2144–2151, 2011.
- [31] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *NIPS*, pages 1097–1105, 2012.
- [32] A. Krull, E. Brachmann, S. Nowozin, F. Michel, J. Shotton, and C. Rother. Poseagent: Budget-constrained 6d object pose estimation via reinforcement learning. In *CVPR*, July 2017.
- [33] A. Kumar and R. Chellappa. Disentangling 3d pose in a dendritic cnn for unconstrained 2d face alignment. *CVPR*, 2018.
- [34] N. Kumar, P. N. Belhumeur, and S. K. Nayar. Facetracer: A search engine for large collections of images with faces. In *ECCV*, pages 340–353, 2008.
- [35] V. Le, J. Brandt, Z. Lin, L. D. Bourdev, and T. S. Huang. Interactive facial feature localization. In *ECCV*, pages 679–692, 2012.
- [36] Y. LeCun, Y. Bengio, and G. E. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.
- [37] Y. Li. Deep reinforcement learning: An overview. *CoRR*, abs/1701.07274, 2017.
- [38] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra. Continuous control with deep reinforcement learning. *ICLR*, 2016.
- [39] J. Lin, Y. Rao, J. Lu, and J. Zhou. Runtime neural pruning. In *NIPS*, pages 2178–2188, 2017.
- [40] F. Liu, D. Zeng, Q. Zhao, and X. Liu. Joint face alignment and 3d face reconstruction. In *ECCV*, pages 545–560, 2016.
- [41] H. Liu, J. Lu, J. Feng, and J. Zhou. Learning deep sharable and structural detectors for face alignment. *TIP*, 26(4):1666–1678, 2017.
- [42] H. Liu, J. Lu, J. Feng, and J. Zhou. Two-stream transformer networks for video-based face alignment. In *TPAMI, in Press*, 2017.
- [43] S. Liu, Z. Zhu, N. Ye, S. Guadarrama, and K. Murphy. Improved image captioning via policy gradient optimization of spider. In *ICCV*, Oct 2017.
- [44] J. Lv, X. Shao, J. Xing, C. Cheng, and X. Zhou. A deep regression architecture with two-stage re-initialization for high performance facial landmark detection. In *CVPR*, July 2017.
- [45] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. A. Riedmiller, A. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.
- [46] A. Newell, K. Yang, and J. Deng. Stacked hourglass networks for human pose estimation. In *ECCV*, pages 483–499, 2016.
- [47] G. Palmer, K. Tuyls, D. Bloembergen, and R. Savani. Lenient multi-agent deep reinforcement learning. *CoRR*, abs/1707.04402, 2017.
- [48] X. Peng, R. S. Feris, X. Wang, and D. N. Metaxas. A recurrent encoder-decoder network for sequential face alignment. In *ECCV*, pages 38–56, 2016.
- [49] P. H. O. Pinheiro and R. Collobert. Recurrent convolutional neural networks for scene labeling. In *ICML*, pages 82–90, 2014.
- [50] A. Pirinen and C. Sminchisescu. Deep reinforcement learning of region proposal networks for object detection. In *CVPR*, June 2018.
- [51] Y. Rao, D. Lin, J. Lu, and J. Zhou. Learning globally optimized object detector via policy gradient. In *CVPR*, June 2018.
- [52] Y. Rao, J. Lu, and J. Zhou. Attention-aware deep reinforcement learning for video face recognition. In *CVPR*, Oct 2017.
- [53] S. Ren, X. Cao, Y. Wei, and J. Sun. Face alignment at 3000 FPS via regressing local binary features. In *CVPR*, pages 1685–1692, 2014.
- [54] Z. Ren, X. Wang, N. Zhang, X. Lv, and L.-J. Li. Deep reinforcement learning-based image captioning with embedding reward. In *CVPR*, July 2017.
- [55] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 300 faces in-the-wild challenge: database and results. *IVC*, 47:3–18, 2016.
- [56] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 300 faces in-the-wild challenge: The first facial landmark localization challenge. In *CVPR*, 2017.

- ICCVW*, 2013.
- [57] E. Sánchez-Lozano, B. Martínez, G. Tzimiropoulos, and M. F. Valstar. Cascaded continuous regression for real-time incremental face tracking. In *ECCV*, pages 645–661, 2016.
- [58] T. Schaul, J. Quan, I. Antonoglou, and D. Silver. Prioritized experience replay. In *ICLR*, 2016.
- [59] J. Shen, S. Zafeiriou, G. G. Chrysos, J. Kossaifi, G. Tzimiropoulos, and M. Pantic. The first facial landmark tracking in-the-wild challenge: Benchmark and results. In *ICCVW*, pages 1003–1011, 2015.
- [60] B. Shi, X. Bai, W. Liu, and J. Wang. Deep regression for face alignment. *arXiv:1409.5230*, 2014.
- [61] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. P. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.
- [62] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. A. Riedmiller. Deterministic policy gradient algorithms. In *ICML*, pages 387–395, 2014.
- [63] G. Song, H. Myeong, and K. Mu Lee. Seednet: Automatic seed generation with deep reinforcement learning for robust interactive segmentation. In *CVPR*, June 2018.
- [64] Y. Sun, X. Wang, and X. Tang. Deep convolutional network cascade for facial point detection. In *CVPR*, pages 3476–3483, 2013.
- [65] J. Supancic, III and D. Ramanan. Tracking as online decision-making: Learning a policy from streaming videos with reinforcement learning. In *ICCV*, Oct 2017.
- [66] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. Deepface: Closing the gap to human-level performance in face verification. In *CVPR*, pages 1701–1708, 2014.
- [67] Y. Tang, Y. Tian, J. Lu, P. Li, and J. Zhou. Deep progressive reinforcement learning for skeleton-based action recognition. In *CVPR*, June 2018.
- [68] G. Trigeorgis, P. Snape, M. A. Nicolaou, E. Antonakos, and S. Zafeiriou. Mnemonic descent method: A recurrent process applied for end-to-end face alignment. In *CVPR*, pages 4177–4187, 2016.
- [69] G. Tzimiropoulos. Project-out cascaded regression with an application to face alignment. In *CVPR*, pages 3659–3667, 2015.
- [70] G. Tzimiropoulos and M. Pantic. Gauss-newton deformable part models for face alignment in-the-wild. In *CVPR*, pages 1851–1858, 2014.
- [71] H. van Hasselt, A. Guez, and D. Silver. Deep reinforcement learning with double q-learning. In *AAAI*, pages 2094–2100, 2016.
- [72] Z. Wang, T. Schaul, M. Hessel, H. van Hasselt, M. Lanctot, and N. de Freitas. Dueling network architectures for deep reinforcement learning. In *ICML*, pages 1995–2003, 2016.
- [73] R. J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8:229–256, 1992.
- [74] W. Wu, C. Qian, S. Yang, Q. Wang, Y. Cai, and Q. Zhou. Look at boundary: A boundary-aware face alignment algorithm. In *CVPR*, 2018.
- [75] Y. Wu and Q. Ji. Discriminative deep face shape model for facial point detection. *IJCV*, 113(1):37–53, 2015.
- [76] Y. Wu and Q. Ji. Robust facial landmark detection under significant head poses and occlusion. In *ICCV*, pages 3658–3666, 2015.
- [77] S. Xiao, J. Feng, J. Xing, H. Lai, S. Yan, and A. A. Kassim. Robust facial landmark detection via recurrent attentive-refinement networks. In *ECCV*, pages 57–72, 2016.
- [78] X. Xiong and F. D. la Torre. Supervised descent method and its applications to face alignment. In *CVPR*, pages 532–539, 2013.
- [79] X. Xiong and F. D. la Torre. Global supervised descent method. In *CVPR*, pages 2664–2673, 2015.
- [80] H. Yang, X. He, X. Jia, and I. Patras. Robust face alignment under occlusion via regional predictive power estimation. *TIP*, 24(8):2393–2403, 2015.
- [81] X. Yu, F. Zhou, and M. Chandraker. Deep deformation network for object landmark localization. In *ECCV*, pages 52–70, 2016.
- [82] S. Yun, J. Choi, Y. Yoo, K. Yun, and J. Y. Choi. Action-decision networks for visual tracking with deep reinforcement learning. In *CVPR*, 2017.
- [83] J. Zhang, S. Shan, M. Kan, and X. Chen. Coarse-to-fine auto-encoder networks (CFAN) for real-time face alignment. In *ECCV*, pages 1–16, 2014.
- [84] Z. Zhang, P. Luo, C. C. Loy, and X. Tang. Learning deep representation for face alignment with auxiliary attributes. *TPAMI*, 38(5):918–930, 2016.
- [85] S. Zhu, C. Li, C. C. Loy, and X. Tang. Face alignment by coarse-to-fine shape searching. In *CVPR*, pages 4998–5006, 2015.
- [86] S. Zhu, C. Li, C. C. Loy, and X. Tang. Unconstrained face alignment via cascaded compositional learning. In *CVPR*, pages 3409–3417, 2016.
- [87] X. Zhu and D. Ramanan. Face detection, pose estimation, and landmark localization in the wild. In *CVPR*, pages 2879–2886, 2012.



Hao Liu received the B.S. degree in software engineering from Sichuan University, China, in 2011, and the Master Engineering degree in computer technology from the University of Chinese Academy of Sciences, China, in 2014, and the Ph.D. degree in Control Science and Technology in Tsinghua University, China, 2018. He is currently an assistant professor with the School of Information Engineering, Ningxia University, China. His research interests include facial analysis and deep learning. He serves as a reviewer for many journals such as IEEE T-PAMI, T-CSVT, Neurocomputing and international conferences such as AAAI, ICME and ICIP.



Jiwen Lu received the Ph.D. degree in electrical engineering from Nanyang Technological University, Singapore, in 2012. He is currently an Associate Professor with the Department of Automation, Tsinghua University, Beijing, China. His current research interests include computer vision, pattern recognition, and machine learning. He is an Associate Editor of the *IEEE Transactions on Image Processing*, the *IEEE Transactions and Circuits and Systems for Video Technology*, the *IEEE Transactions on Biometrics, Behavior, and Identity Science, Pattern Recognition*, and the *Journal of Visual Communication and Image Representation*.



Minghao Guo is currently pursuing the B.S. degree at the department of automation, Tsinghua University. He will pursue the Ph.D. degree at the department of automation, Tsinghua University in September, 2018. His research interests include face alignment, face detection and deep learning.



Suping Wu received the B.S. degree from the school of mathematics and computer from Ningxia University, China, in 1986, and the M.S. degree in computer software and theory from Beijing Normal University, China, in 2002. She is currently a professor with the School of Information Engineering, Ningxia University, China. Her research interests include big data analysis, parallel distributed processing and visual 3D reconstruction. She also serves as an associate editor of the Journal on Communications.



Jie Zhou received the B.S. and M.S. degrees from the Department of Mathematics, Nankai University, Tianjin, China, in 1990 and 1992, respectively, and the Ph.D. degree from the Institute of Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan, China, in 1995. Since 2003, he has been a Full Professor with the Department of Automation, Tsinghua University. His current research interests include computer vision, pattern recognition, and image processing. He is an Associate Editor of the *IEEE Transactions and Pattern Analysis and Machine Intelligence*.