



HAL
open science

Reinforcement Learning coupled with Finite Element Modeling for Facial Motion Learning

Duc-Phong Nguyen, Marie-Christine Ho Ba Tho, Tien-Tuan Dao

► **To cite this version:**

Duc-Phong Nguyen, Marie-Christine Ho Ba Tho, Tien-Tuan Dao. Reinforcement Learning coupled with Finite Element Modeling for Facial Motion Learning. *Computer Methods and Programs in Biomedicine*, 2022, pp.106904. 10.1016/j.cmpb.2022.106904 . hal-03676366

HAL Id: hal-03676366

<https://hal.science/hal-03676366v1>

Submitted on 22 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **Reinforcement Learning coupled with Finite Element Modeling**
2 **for Facial Motion Learning**

3 Duc-Phong NGUYEN¹, Marie-Christine HO BA THO¹, Tien-Tuan DAO²

4 ¹ Université de technologie de Compiègne, CNRS, Biomechanics and Bioengineering,
5 Centre de recherche Royallieu, CS 60 319 - 60 203, Compiègne Cedex, France

6 ² Univ. Lille, CNRS, Centrale Lille, UMR 9013 - LaMcube - Laboratoire de Mécanique,
7 Multiphysique, Multiéchelle, F-59000 Lille, France

8 duc-phong.nguyen@utc.fr, hobatho@utc.fr, tien-tuan.dao@centralelille.fr

9

10 Manuscript submitted as a **Research Paper** to the

11 *Computer Methods and Programs in Biomedicine*

12 (2nd Revision) May 2022

13 Corresponding author: Prof. Tien Tuan Dao

14 Centrale Lille Institut, CNRS UMR 9013 - LaMcube

15 Laboratoire de Mécanique, Multiphysique, Multiéchelle

16 59655 Villeneuve d'Ascq Cedex, France

17 Tel: 33 3 20 43 43 04

18 E-mail: tien-tuan.dao@centralelille.fr

19

20 **Abstract**

21 **Background and Objective:** Facial palsy patients or patients with facial transplantation have abnormal
22 facial motion due to altered facial muscle functions and nerve damage. Computer-aided system and physics-
23 based models have been developed to provide objective and quantitative information. However, the
24 predictive capacity of these solutions is still limited to explore the facial motion patterns with emerging
25 properties. The present study aims to couple the reinforcement learning and the finite element modeling for
26 facial motion learning and prediction.

27 **Methods:** A novel modeling workflow for learning facial motion was developed. A physically-based model
28 of the face within the Artisynth modeling platform was used. Information exchange protocol was proposed to
29 link reinforcement learning and rigid multi-bodies dynamics outcomes. Two reinforcement learning
30 algorithms (deep deterministic policy gradient (DDPG) and Twin-delayed DDPG (TD3)) were used and
31 implemented to drive the simulations of symmetry-oriented and smile movements. Numerical outcomes were
32 compared to experimental observations (Bosphorus database) for evaluation and validation purposes.

33 **Results:** As result, after more than 100 episodes of exploring the environment, the agent starts to learn from
34 previous trials and can find the optimal policy after more than 300 episodes of training. Regarding the
35 symmetry-oriented motion, the muscle excitations predicted by the trained agent help to increase the value of
36 reward from $R = -2.06$ to $R = -0.23$, which counts for ~89% improvement of the symmetry value of the face.
37 For smile-oriented motion, two points at the edge of the mouth move up 0.35 cm, which is within the range
38 of movements estimated from the Bosphorus database (0.4 ± 0.32 cm).

39 **Conclusions:** The present study explored the muscle excitation patterns by coupling reinforcement learning
40 with a detailed finite element model of the face. We developed, for the first time, a novel coupling scheme to
41 integrate the finite element simulation into the reinforcement learning for facial motion learning. As
42 perspectives, this present workflow will be applied for facial palsy and facial transplantation patients to guide
43 and optimize the functional rehabilitation program.

44 **Keywords:** Reinforcement learning, finite element modeling, facial motion learning, facial rehabilitation,
45 Artisynth.

46 **1. Introduction**

47 Facial palsy patients or patients with facial transplantation have abnormal facial motion
48 patterns due to altered facial muscle functions and nerve damage leading to abnormal
49 motion control for different movements such as eating, speaking, or facial expressions [1-
50 4]. Moreover, involved patients also suffer asymmetric face effect, which indicates the
51 imbalance and inequality of facial structure in terms of shape, size, location, and
52 arrangement of left and right components on the sagittal plane [38]. In fact, the recovery of
53 a symmetric face with balanced functionalities requires a complex rehabilitation process in
54 which patients must practice patient specific facial movements. Thus, understanding of
55 facial motion mechanism helps the involved patients to recover symmetrical movements
56 and normal facial expressions. It is important to note that current facial rehabilitation has
57 mainly based on a mirror approach to monitor the visual qualitative feedback from the
58 rehabilitation exercise. More precisely, patients watch their distorted features in the mirror
59 as a reference to teach themselves the right expressions during rehabilitation exercises.
60 This strategy is ineffective and subjective without any feedback. Moreover, the current
61 rehabilitation process is limited by a lack of patient specific knowledge about muscles
62 driving facial motions. Therefore, understanding of facial motion mechanism and muscle
63 activation and coordination is clearly fundamental. To provide quantitative and objective
64 information on the facial motion during the rehabilitation exercise, computer-based
65 systems that automatically recognizes action units (AUs) defined by the Facial Action
66 Coding System (FACS) have been developed [5]. Such complex systems can provide an
67 objective guideline for monitoring the facial rehabilitation process, which is a long-term,
68 inconvenient, and sometimes ineffective [6-8].

69 In addition, for investigating muscles driving facial motion problem, biomechanical
70 models are recommended because they can be customized to reflect the true anatomy and
71 pathological anatomical deformations as well as imitate physical process [39]. In fact,
72 physics-based facial models using finite element methods have been intensively developed
73 to explore the role of the facial muscle excitation, contraction and coordination during
74 facial motion [9-16, 31-32]. Muscle excitation represents the neural control process, which
75 contracts the face tissues and moves the skull to perform facial expressions and movements.
76 Despite a detailed view on the muscle contraction mechanism and its effect on the facial
77 motion, the physics-based approach is descriptive with a priori known input information
78 such as muscle properties. Moreover, which muscles and what value of muscle excitations
79 for performing a desired movement for facial rehabilitation is still an open and
80 longstanding research question. It is practically impossible to directly measure muscle
81 activations from living subjects due to safety and accessibility limitations. Diverse
82 numerical techniques have been proposed for estimating muscle excitation such as inverse
83 dynamics, forward-dynamics tracking simulation, and optimal control strategies [10].
84 However, the use of this approach depends strongly on the a priori definition of input data,
85 model properties and the targeted motion. Thus, this approach has a limited predictive
86 capacity to explore a larger parameter space to find emerging properties during dynamic
87 movements of the face.

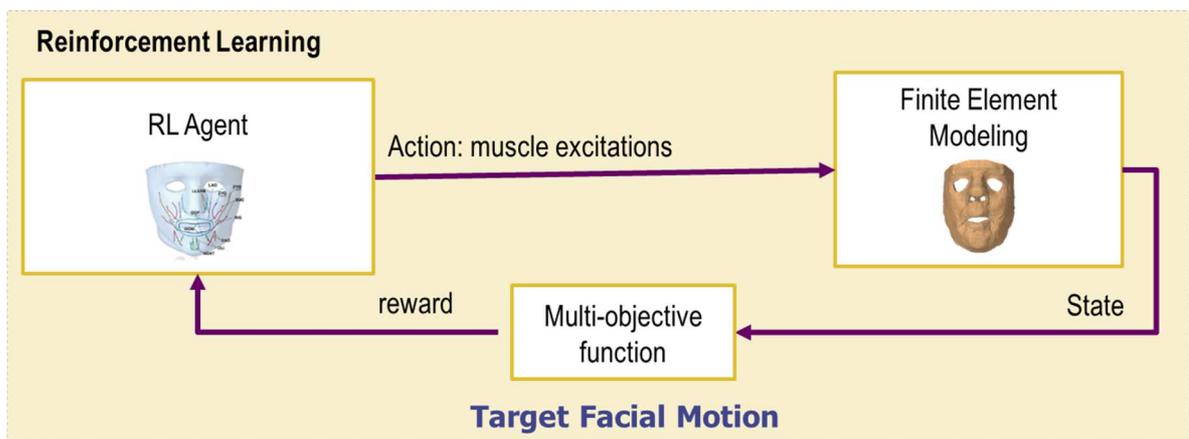
88 In spite of the increasing availability of massive databases and computational models,
89 artificial intelligence has rapidly grown [17]. One of the promising solutions in the control
90 field is the reinforcement learning with tremendous theoretical and practical achievements
91 in robotics control [18], gaming [19], autonomous driving [20], computer vision [21],
92 and healthcare [22-24]. In particular, the question of the use of this learning strategy in the

93 healthcare domain to tackle real-world applications has recently raised [22-24].
94 Reinforcement learning distinguishes from other types of machine learning in several
95 perspectives. The agent collects data through interactions with the environment and uses
96 that data to train the agent itself. This dependence results in variation outcomes from one
97 run to another. Recently, the reinforcement learning strategy has been coupled with rigid
98 multi-bodies dynamics to explore the motion of the lower limbs during walking and age-
99 related falls [33]. Thus, this learning strategy opens new avenues to explore human system
100 motion and novel emerging properties without any a priori motion data. Thus, the present
101 study aims to explore the facial motion learning capacity by the coupling between the
102 reinforcement learning and the finite element modeling. The main objective is to provide,
103 for the first time, the modeling workflow for this complex coupling and then to evaluate
104 different learning strategies to establish motion patterns of the face during facial expression
105 motions. Our novel solution will explore the patient specific facial motions without a priori
106 data from the patient and then provides a set of facial muscle activation and coordination
107 patterns for a specific rehabilitation-oriented movement (e.g. symmetry or smile). The
108 remainder of the paper is organized as follows: section 2 focuses on the coupling workflow
109 between reinforcement learning and finite element model of the face for learning symmetry
110 and smile motions. Section 3 provides computational results and the comparison with other
111 studies. Section 4 provides a detail discussion on the method and obtained results. Finally,
112 section 5 addresses conclusions and perspectives of the present work.

113 **2. Materials and methods**

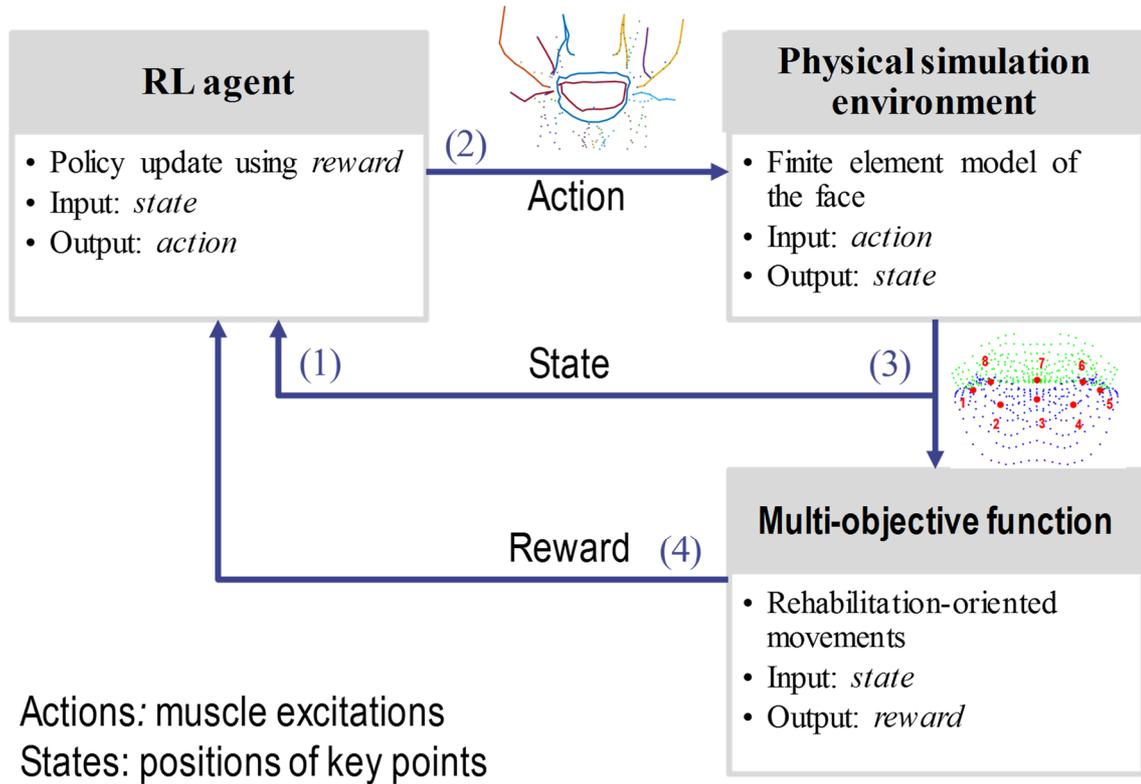
114 **2.1 Novel coupling workflow between reinforcement learning and finite element** 115 **modeling**

116 Our novel simulation workflow requires main two components (Fig. 1): 1) A
117 reinforcement learning agent (a human face) having a policy that decides what action
118 (muscle excitations) to take when it observes a state (facial motion) and 2) A finite element
119 modeling and simulation environment. The coupling between the finite element simulation
120 environment and the reinforcement learning process is managed by an information
121 exchange protocol. More precisely, at the beginning, the reinforcement learning agent
122 observes the state of the face using the positions of selected key points (Fig. 2). Secondly,
123 the policy predicts values of muscle excitations, which are then applied to the
124 biomechanical model of the face for a physical simulation. Then, the simulation
125 environment returns the positions of selected key points after simulation. And finally, these
126 positions are used to compute the reward value by pre-designed multi-objective function
127 (related to symmetry or smile exercises), which is then used to update training parameters
128 for the training process.



129

130 *Figure 1. Overview of the novel coupling workflow between reinforcement learning and*
131 *finite element modeling.*



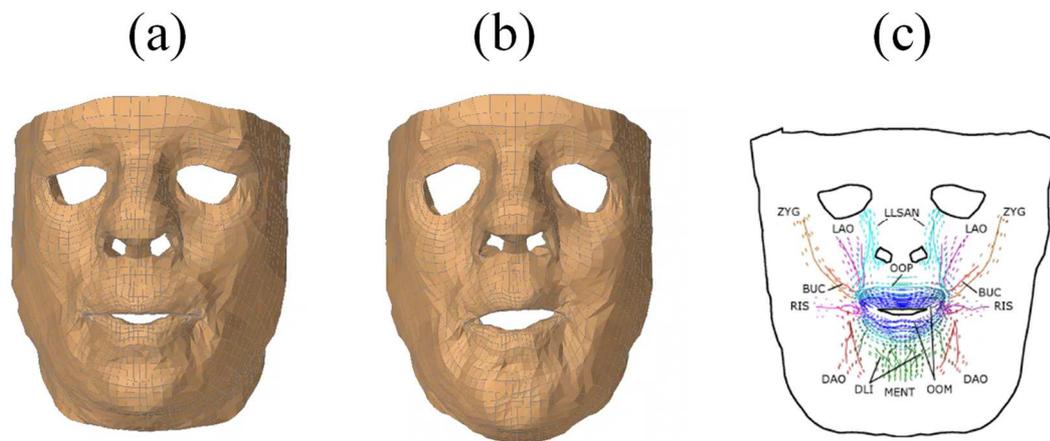
132

133 *Figure 2. Detailed flowchart of the interaction between reinforcement learning and finite*
 134 *element modeling processes.*

135 **2.2 Face finite element model**

136 A physically-based model of the face within the Artisynt modeling platform was used
 137 (Fig. 3a). This model has been from previous researches [16], [25], [26], [27]. The face
 138 finite element model includes three components such as 1) soft-tissue component with the
 139 hypodermis, dermis, and epidermis layers, 2) a cranium and maxilla component, 3) a jaw-
 140 hyoid component [28]. To reduce computational cost and accelerate the training process,
 141 the facial model is simplified by keeping only the soft-tissue component with ten orofacial
 142 muscles (Levator Anguli Oris (LAO), Levator Labii Superioris Alaeque Nasi (LLSAN),
 143 Buccinator (BUC), Zygomaticus (ZYG), Depressor Anguli Oris (DAO), Risorius (RIS),
 144 Depressor Labii Inferioris (DLI), Mentalis (MENT), Orbicularis Oris Peripheralis (OOP),

145 Orbicularis Oris Marginalis (OOM)) (Fig. 3c). The soft tissue finite element mesh consists
146 of 6342 brick elements (with 6024 hexahedrons and 318 wedges) and 8720 nodes. The
147 activation for the face model results from the orofacial muscle strain and force.
148 Ten orofacial muscles are modeled and attached in the lower face that applies muscle
149 forces in terms of muscle excitations onto the finite element model. Muscle fibers are
150 modeled by a set of uniaxial cable elements. For example, the zygomatic ligaments are
151 represented by fixing all degrees-of-freedom of soft tissue nodes that are in the region
152 where these ligaments attach to the maxilla. Soft tissue constitutive equation for the
153 hypodermis layer is based on a Mooney-Rivlin constitutive equation, and Fung constitutive
154 equation for the epidermis and dermis layer as in the Flynn et al. paper [16]. The
155 mechanical characteristics (such as force-displacement response, pre-stress behaviors, non-
156 linear, anisotropic, and viscoelastic constitutive laws) for the skin layer were estimated
157 based on a combination of in vivo experiments and numerical methods. Muscles are
158 modeled as continuous sets of cable elements, which activate in tension as point-to-point
159 Hill-type models and are aligned along element edges. The mechanical property evolution
160 of muscle contraction comprises muscle contractile fibres (active part), muscle body
161 (passive part), and the stress stiffening effect [25]. The movements of the mandible
162 generated by muscles of mastication are not handled yet in the model. Thus, the superficial
163 muscles, which are muscles around the lip region, involved in facial mimics are focused.
164 Two finite element models of the face corresponding with the modeling of the symmetric
165 face (Fig. 3a) and the asymmetric face (Fig. 3b).



166

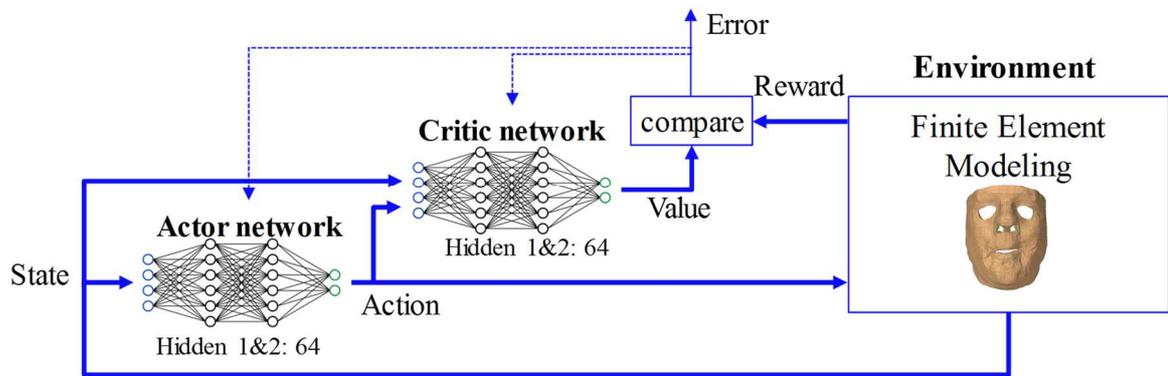
167 *Figure 3. The face finite element model (a) referred as the symmetric face, (b) referred as*
 168 *the asymmetric face (unbalanced deformation between left and right sides), and related*
 169 *facial muscle network (c).*

170 **2.3 Reinforcement Learning for Facial Motion Control**

171 **2.3.1 Reinforcement learning model and algorithms**

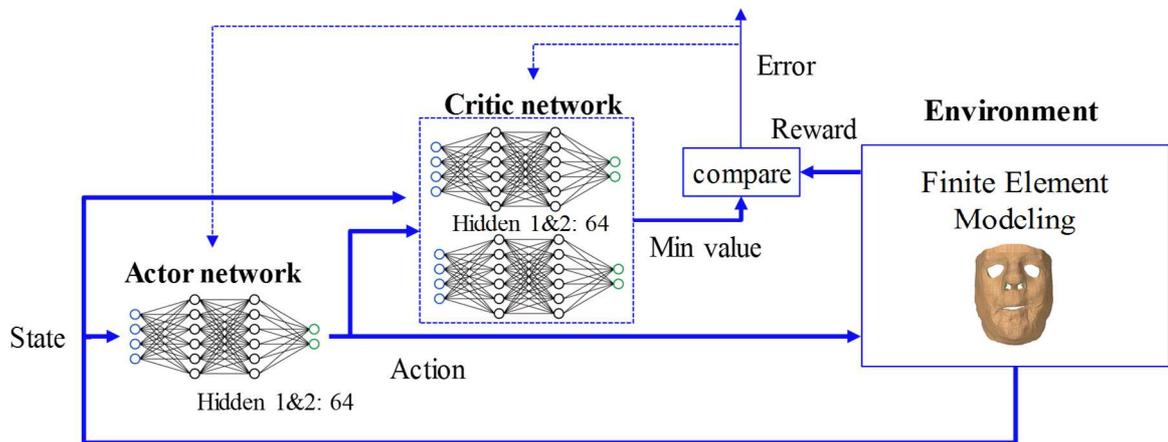
172 Reinforcement learning (RL) aims to find a policy, $\pi(a|s)$, which maps the state space to
 173 the action space and instructs the agent on how to make decisions that maximizes the long-
 174 term cumulative reward inspired by a reward function $r(s, a)$, where a is the action needs
 175 to take in the state s . Bellman equations are solved to find the optimal policy. In this
 176 present study, two RL algorithms were used. The first algorithm is the Deep Deterministic
 177 Policy Gradient (DDPG) in which the Bellman equation was solved by combining a deep
 178 neural network for learning Q function and a deterministic policy gradient algorithm for
 179 learning a policy. This is off-policy reinforcement learning used for continuous state and
 180 action spaces, which is suitable for our problem. The second used algorithm is the Twin
 181 Delayed DDPG (TD3). DDPG is often brittle with the tuning process for hyperparameters.
 182 It usually fails when exploiting the error in the Q-function, the learned Q function starts to

183 overestimate Q-values results in policy breaking. Twin delayed DDPG copes with this
 184 issue and improves performance by applying three tricks as 1) Clipped Double-Q Learning:
 185 learning two Q-functions (twin) and using smaller Q-values for the Bellman error loss
 186 functions, 2) Delayed Policy Updates: sparse updating the policy compared to the Q-
 187 function, 3) Target Policy Smoothing: adding noise to target actions to reduce exploiting
 188 Q-function errors.



189

190 Figure 4. The network architecture of DDPG: one actor network and one critic network.



191

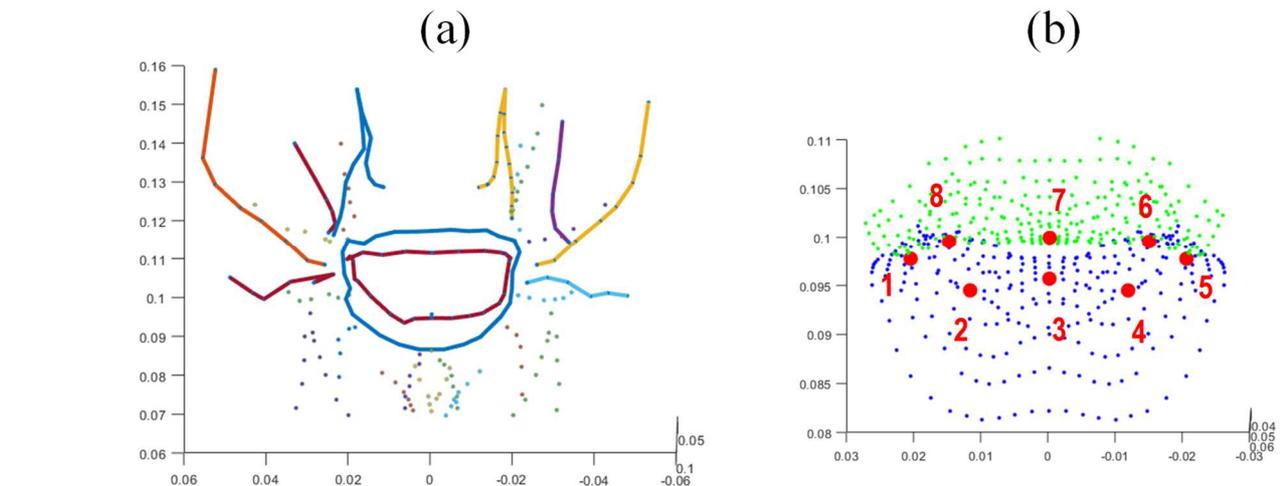
192 Figure 5. The network architecture of TD3: one actor network and two critic networks.

193 The network architecture of DDPG contains the actor network and the critic network. Each
 194 network has two hidden layers with 64 nodes (Fig. 4). The actor network inputs the state
 195 vector while outputs the action vector. The input of the critic network contains both the

196 action vector output from the actor network and the state vector, while the output is the
197 predicted Q-value. TD3 has the same architecture as in DDPG except it has two critic
198 networks (Fig. 5).

199 2.3.2 Reward Function, Action Space, and State Space

200 The aim of our study is to find the appropriate muscle excitations for performing a facial
201 motion, which is generated by defining appropriate biomechanics-inspired reward function.
202 In our model, action is a vector of 10 pairs of left and right muscles in terms of muscle
203 forces normalized between 0 and 1. To avoid the exhausted search, only significant
204 muscles (*left and right Levator Anguli Oris (LAO)*, *left and right Levator Labii Superioris*
205 *Alaeque Nasi (LLSAN)*, *left and right Zygomatics (ZYG)*, *left and right Risorious (RIS)*,
206 *Orbicularis Oris Marginalis (OOM)*, *Orbicularis Oris Peripheralis (OOP)*) were included
207 in training process (Fig. 6a). In our model, the agent's state was defined through a set of
208 landmark points focusing on the mouth region of the face. In fact, 8 key points on the lips
209 are chosen as representations of the state of the face (Fig. 6b).



210

211 *Figure 6. Selected muscles excitations for training (a) and landmark points for the RL*

212

agent's state (b).

213 Regarding the reward function, the agent receives a reward value from environment at each
 214 time-step. Note that the training efficiency of reinforcement learning algorithm depends
 215 strongly on defining the reward function. In our study, the reward function is designed by a
 216 motion-oriented (e.g. symmetry-targeted motion, smile expression, sound pronunciations)
 217 strategy. More precisely, different reward functions were formulated using the Euclidean
 218 distance and angle created from the defined 8 landmark points. Mathematically, reward
 219 functions are defined as follows:

$$220 \quad R_{symmetry}^{distance} = -1000 * (r_1^d + r_2^d + r_3^d) \quad (1)$$

$$221 \quad R_{symmetry}^{angle} = -(r_1^a + r_2^a + r_3^a) \quad (2)$$

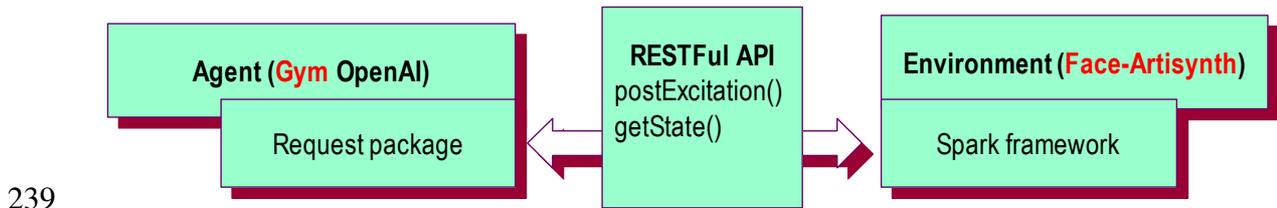
$$222 \quad R_{smile} = \Delta d \quad (3)$$

223 where r_i^d, r_i^a are the symmetry value-based distance and angle between left side compared
 224 to right side for each pair point (point 1-5, 2-4, 8-6). Δd is the total moving up of point #1
 225 and point #5 defined in Fig. 6b.

226 **2.4 Information exchange protocol and implementation**

227 Our study is based on two modeling platforms (i.e. Artisynt and PyTorch) coming from
 228 different fields. The exchange information between these platforms need a novel
 229 communication protocol. Artisynt-RL has been proposed to open the exchange capacity
 230 with rigid multi-bodies dynamics simulation [30]. In the present study, Artisynt-RL was
 231 extended to exchange information between the PyTorch platform which is a Python-based
 232 training platform for RL models and Artisynt for running face finite element simulation.
 233 Note that Artisynt-RL is a cross-platform based java script. To achieve this objective,
 234 different technologies (RESTful API as a plugin, Spark framework from java, and Request

235 package from python) were used as shown in Fig. 7. Regarding the communication
236 protocol, the muscle excitations from the reinforcement learning module are posted into
237 the Artisyng module, then a new state is obtained from the Artisyng module to
238 reinforcement learning module after each simulation step.



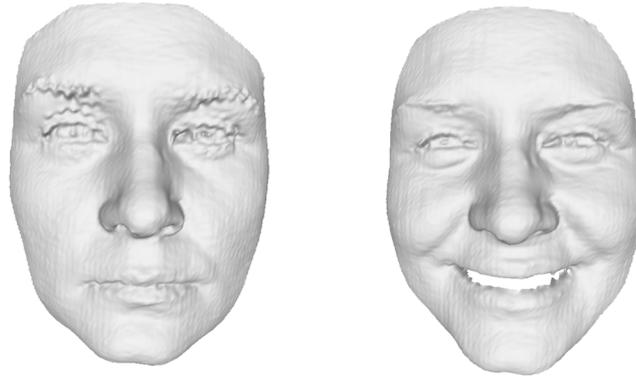
239
240 *Figure 7. RESTful API as a plugin for bridging reinforcement learning and Artisyng*

241 As hardware configuration, a virtual machine configuration with ubuntu 20.04, 8 CPU, 16
242 Gb RAM, Python 3.6, and the open-source stable-baselines3 was used for the training
243 process.

244 **2.5 Evaluation and validation**

245 An open access 3D face database, named Bosphorus, was used for evaluation and
246 validation purpose. This database includes 105 subjects (44 females and 61 males) with
247 different expressions, poses, and occlusion conditions. 65 subjects have 7 expressions such
248 as happiness (smile), surprise, fear, sadness, anger, disgust, and neutral. Happiness (smile)
249 and neutral expressions, which are available in 130 face scans of all the subjects, were used
250 for further validation. Firstly, three key point at positions #1, #5, and #7 (as in Fig. 6b)
251 were manually picked for each face scan. Secondly, all the face scans were transferred
252 such that the point at position #7 is at the origin (coordinate [0, 0, 0]) and face scans of the
253 same person is at the same orientation. Finally, the total displacement of moving up action
254 of the two key points at positions #1 and #5 (as in Fig. 6b) were computed by subtracting
255 the corresponding points of smile face scan and neutral face scan of the same person. In

256 fact, the values of the used reward functions were computed for each posture (neutral and
257 smile expression). Obtained values were represented in mean and standard deviation and
258 then compared to the final outcomes from the RL process.



259

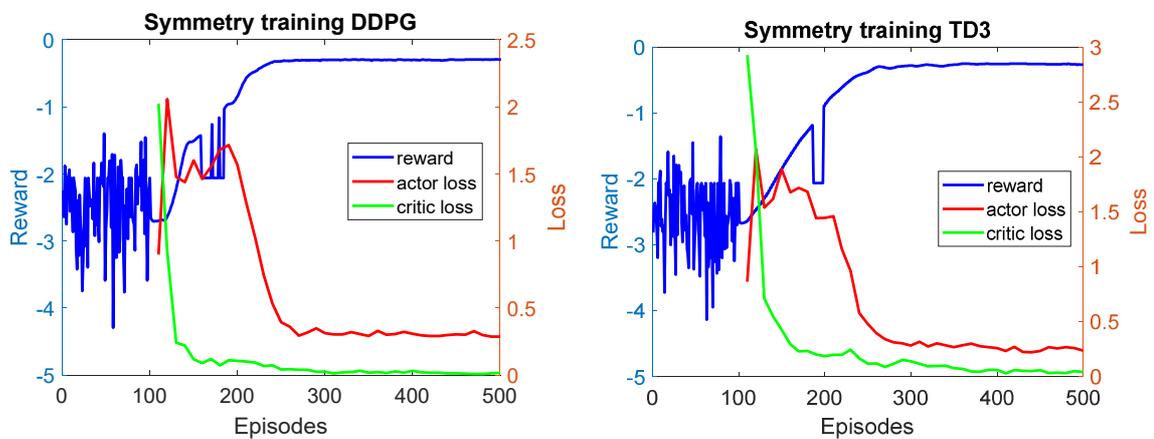
260 *Figure 8. Illustration from Bosphorus database with two expressions: neutral (left) and*
261 *smile (right).*

262 A hyperparameter tuning process was implemented to select the best neural network
263 architecture and parameters. In particular, the hyperparameter tuning process is not
264 automatically tuned, but manually selects each parameter, while other parameters remain
265 unchanged. Each trial was performed for a training task with the hyperparameter within a
266 predefined set to ensure that the agent successfully explores and learns to make decisions
267 in its environment. The predefined set of hyperparameters includes several most critical
268 parameters, which govern the performance of reinforcement learning such as the neural
269 network size (nodes in hidden layers ([64, 64] or [400, 300] for the actor and the critic
270 networks)), learning rate (0.001, 0.01, note that the learning rate is shared for all networks),
271 batch size (16, 32), τ parameter, which used to soft update both critic and actor target
272 networks (0.001, 0.005).

273 **3. Computational results**

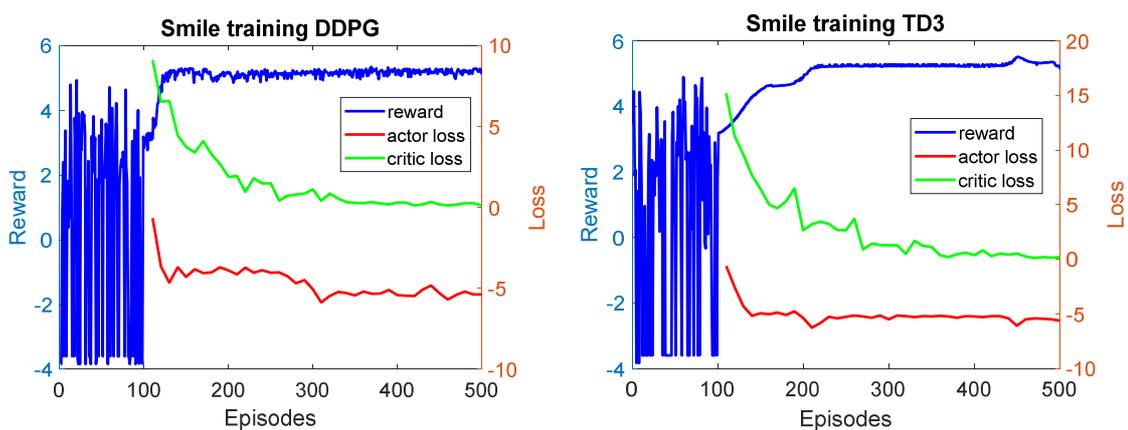
274 3.1 RL accuracy and performance

275 Figure 9 shows the reward and the loss evolutions during the training process of the agent
276 to conduct one symmetry action. In general, after more than 100 episodes of random
277 interaction in the environment, the agent starts to learn from previous trials and can find
278 the optimal policy after more than 300 episodes of training. In particular, the learning start
279 parameter was set to 100 in the training phase, allowing the reinforcement learning agent
280 to collect a set of transitions ($\mathcal{D} = (s, a, r, s', d)$) by performing a random action from the
281 action space to the environment before learning from previous trials. These random actions
282 result in instable trend in the reward values in first 100 episodes. Having just 200 episodes
283 of learning, the agent still learns and explores the environment to discover the optimal
284 policy. During learning, the not optimal policy may predict the random actions for
285 exploring more the environment that might dramatically drop reward values. From 200 to
286 300 episodes of learning, the agent gradually finds the optimal policy after more than 300
287 episodes of training. The reduction in actor loss and critic loss values demonstrates the
288 efficacy of learning strategy in both DDPG and TD3 methods. The reward value predicted
289 using the model trained by TD3 ($R = -0.26$) is slightly higher than that of using the model
290 trained by DDPG ($R = -0.33$).



292 *Figure 9. The reward value and the loss values of the actor network and the critic network*
 293 *during training reinforcement learning agent with two different methods: DDPG (left),*
 294 *TD3 (right) for symmetry-oriented functional rehabilitation using 4 muscles as ZYG, RIS,*
 295 *OOM, OOP. The reward value predicted using the model trained by DDPG is $R = -0.33$.*
 296 *The reward value predicted using the model trained by TD3 is $R = -0.26$.*

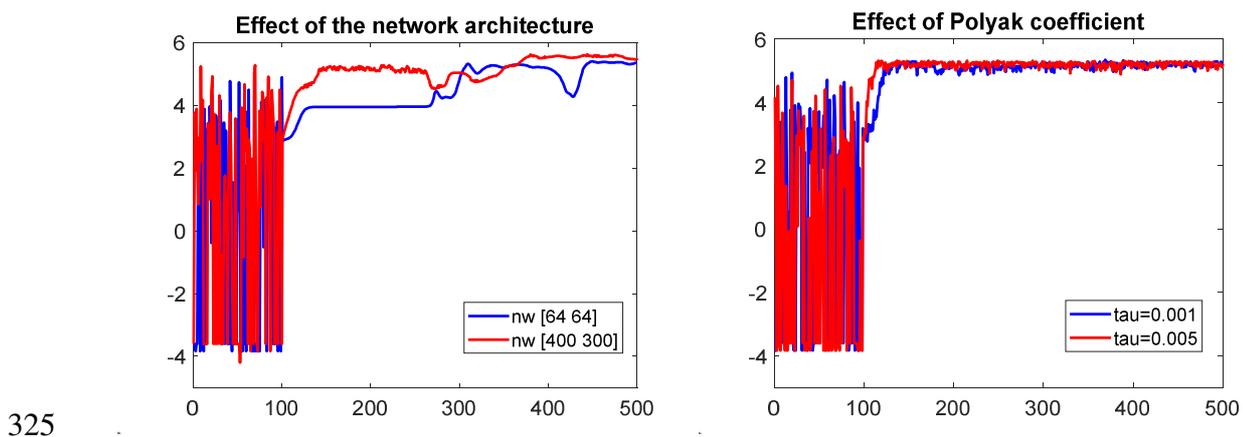
297 Figure 10 shows the reward and the loss evolutions during the training process of the agent
 298 to perform the smile action. The similar patterns are observed according to figure 9. More
 299 precisely, the agent spends 100 episodes collecting a set of transitions by taking random
 300 actions to the environment resulting in instability of reward values. It then starts to learn
 301 from previous trials and can find the optimal policy after more than 200 episodes of
 302 training. The training is successfully demonstrated by the reduction of the loss value of
 303 both actor and critic networks. The reward values during training model using DDPG
 304 algorithm seem noisier compared to that of training model using TD3 algorithm. The
 305 reward value predicted using the model trained by TD3 ($R = 5.36$) is slightly higher than
 306 that of using the model trained by DDPG ($R = 5.26$).



307
 308 *Figure 10. The reward value and the loss values of the actor network and the critic*
 309 *network during training reinforcement learning agent with two different methods: DDPG*

310 (left), TD3 (right) for smile-oriented functional rehabilitation using 3 facial muscles as
311 LAO, LLSAN, ZYG. The reward value predicted using the model trained by DDPG is $R =$
312 5.26. The reward value predicted using the model trained by TD3 is $R = 5.36$.

313 Table 1 shows the reward values obtained during the hyperparameter tuning process for
314 DDPG method when training smile expression. The process can help to identify the better
315 network architecture with associated optimal set of hyperparameters. In fact, the
316 reinforcement learning architecture including two hidden layers [400, 300] for the actor
317 and the critic networks, a batch size of 16, a learning rate for all networks of 0.001, τ
318 parameter of 0.005, and without action noise yields the best reward value. The
319 computational time for training 500 episodes for each smile training and symmetry training
320 is around 6 hours. However, the most computational time is in the simulation environment,
321 where each simulation lasts for 30 seconds (5 hours for 500 episodes related to simulation
322 and restart of Artisynth only). Figure 11 reports the effect of the network architecture and τ
323 parameter. The network architecture has a more important effect than that of the τ
324 parameter.



326 *Figure 11. Reward value during training reinforcement learning agent with different*
327 *hyperparameters and the network architecture.*

Table 1. The reward values obtained during the hyperparameter tuning process

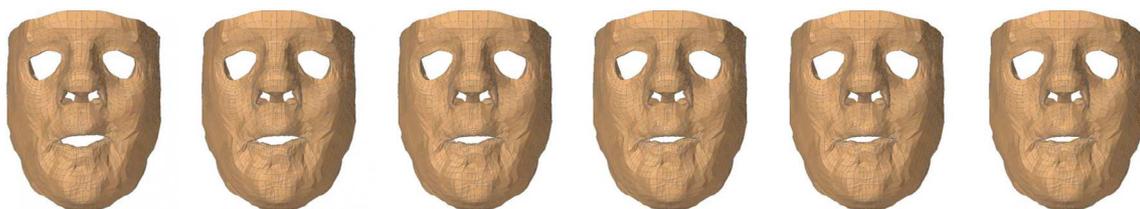
Hyperparameters	Reward
Batch size	16
	32
Learning rate	0.01
	0.001
Network architecture	[64, 64]
	[400, 300]
τ parameter	0.001
	0.005
Action noise	With
	Without

328

329 To sum up, two reinforcement learning algorithms namely DDPG and TD3 have been used
 330 for learning symmetry and smile motions of the finite element model of the face. In terms
 331 of reward predicted from the trained model, TD3 seems to have slightly better performance
 332 compared to DDPG. The hyperparameter tuning process was also proposed to find the
 333 suitable parameter for the model.

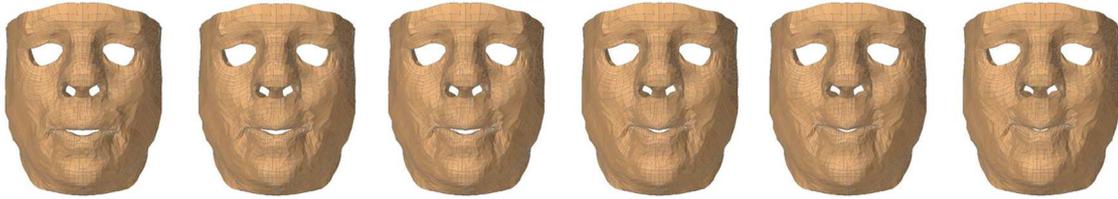
334 **3.2 Facial motion learning**

335 The obtained outcome and corresponding muscle excitation value of the symmetry-
 336 oriented functional rehabilitation is shown in Fig. 12. and Table 2. According to the
 337 prediction, the muscle on the right side of the mouth such as right OOM, right RIS, and
 338 right ZYG are activated, while only the left OOP muscle is activated to improve the
 339 symmetry of the face from the initial state with the reward value $R = -2.06$ to the new state
 340 with the reward value $R = -0.23$, which counts 88.8% improvement.



341

342 *Figure 12. Face animation for symmetry-oriented motion. The face at initial state (on the*
 343 *left $R = -2.06$) and after received muscle excitation (on the right $R = -0.23$) output from*
 344 *reinforcement learning for symmetry-oriented functional rehabilitation.*



345

346 *Figure 13. Face animation for smile-oriented motion. The face at initial state (on the left R*
 347 *= -1.6) and after received muscle excitation (on the right $R = 5.35$) output from*
 348 *reinforcement learning for smile-oriented motion.*

349 Regarding the smile-oriented motion simulation of the finite element model of the face,
 350 both left and right muscles of LAO and ZYG are activated, while LLSAN is not activated
 351 as in Fig. 13 and Table 3. The measured reward value increases from -1.6 at the initial state
 352 to 5.3 at the terminal state. The obtained muscle activation levels for smiling movement are
 353 within the range of values reported by Flynn et al. [16]. However, it is important to note
 354 that there is a difference in smiling patterns between our simulation (i.e. unconstrained
 355 smile) and their simulations (i.e. smiles with open mouth or closed mouth).

356 The muscle action line length change and contraction amplitude $\xi^{CE} = \frac{L-L_0}{L_0} = \frac{\Delta L}{L_0}$ are
 357 shown in Table 2. The contraction amplitudes of OOM and OOP are estimated as the area
 358 that these muscles cover $\xi^{CE} = \frac{S-S_0}{S_0} = \frac{\Delta S}{S_0}$. Related to the smile, the right ZYG contracts -
 359 16.26%, while this number on the left is -15.12%. The right and left LAO contract around -
 360 30%. Note that all muscle contraction levels during smiling are in good agreement with
 361 those estimated using Kinect-driven rigid multi-bodies modeling (Nguyen et al. [40]). In

362 their study [40], the head model was reconstructed from the subject-specific data acquired
363 by a Kinect device in a smile position. Then, a skull model was generated using the
364 statistical fitting method from a generic skull model and the head model. Moreover, a
365 muscle network was defined using insertion points on the head model and attachment
366 points from the skull model. Finally, the muscle contraction levels were estimated using
367 the length of these insertion points and attachment points.

Table 2. Muscle contraction levels during different facial expressions and comparison to the literature data.

Muscle / ξ^{CE}	symmetry		Smile		Nguyen et al. [40] (smile)
	L_0 (mm) / S_0 (mm ²)	$\frac{\Delta L}{L_0}$ (%) / $\frac{\Delta S}{S_0}$ (%)	L_0 (mm) / S_0 (mm ²)	$\frac{\Delta L}{L_0}$ (%) / $\frac{\Delta S}{S_0}$ (%)	$\frac{\Delta L}{L_0}$ (%)
Right ZYG	52 mm	-2.19	54.5 mm	-16.31	From -9.13 to -19.72
Left ZYG	52.2 mm	1.12	54.6 mm	-15.09	From -13.59 to -21.32
Right LLSAN	27.2 mm	-1.62	27.9 mm	-10.26	From -1.99 to -8.12
Left LLSAN	27.2 mm	-0.38	27.9 mm	-10.2	From -0.69 to -6.13
Right LAO	27.3 mm	-1.18	30 mm	-28.13	From -18.66 to -29.46
Left LAO	24.3 mm	1.56	30 mm	-28.17	From -21.19 to -28.03
Right RIS	52.2 mm	-6.21	52.9 mm	-8.12	From 3.55 to 7.30
Left RIS	52 mm	2.44	52.9 mm	-8.135	From -3.09 to 6.96
OOM	590 mm ²	-12.97	665 mm ²	-2.01	-
OOP	1099 mm ²	-7.98	1138 mm ²	17.40	-

368

369

Table 3. Muscle activation levels reported from our simulation and its comparison to the literature data

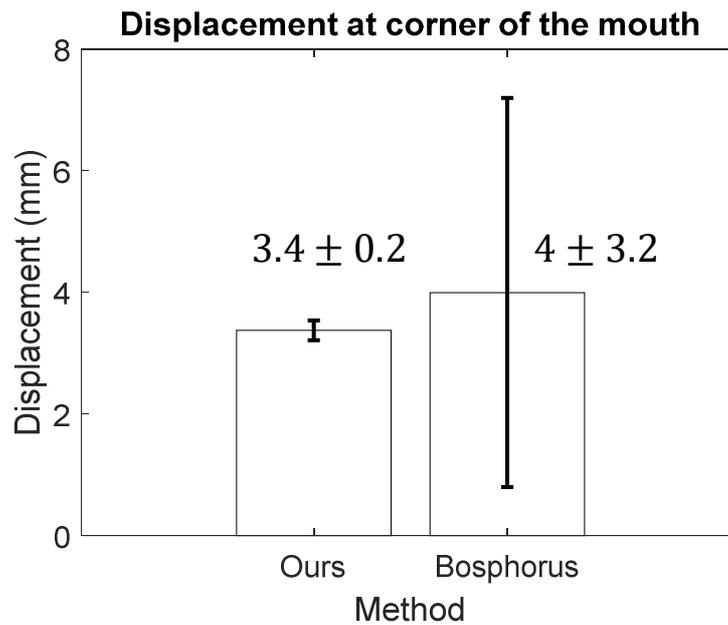
Muscle	symmetry	smile	Flynn et al. [16] (closed mouth smile)	Flynn et al. [16] (open mouth smile)
Right ZYG	0.2	0.4	0.2	0.5
Left ZYG	0	0.4	0.2	0.5
Right LLSAN	0	0	0.1	0.5
Left LLSAN	0	0	0.1	0.5
Right LAO	0	0.4	0.1	0.5
Left LAO	0	0.4	0.1	0.5
Right RIS	0.4	0	0.2	0.6

Left RIS	0	0	0.2	0.6
Right OOM	0	0	0	0
Left OOM	0	0	0	0
Right OOP	0.1	0	0	0
Left OOP	0.4	0	0	0

370

371 **3.3 Evaluation and validation**

372 For symmetry-oriented motion, the muscle excitations predicted by the trained agent help
373 to increase the value of reward from $R = -2.06$ to $R = -0.23$, which counts for $\sim 88.8\%$.
374 While this number for smile-oriented motion, the reward value for both corners of the
375 mouth increases from $R = -1.6$ at the initial state to around $R = 5.3$ at the terminal state,
376 which is 0.35 cm moving up on average for each corner of the mouth. The average distance
377 of moving up for each corner of the mouth was 3.4 ± 0.2 estimated using reward values
378 from Table 1 for different hyperparameters of the reinforcement learning algorithm. This is
379 within the range of movements compared to the value calculated from the Bosphorus
380 database, this value is 0.4 ± 0.32 cm when a person makes maximum effort to smile as in
381 Fig. 14. The variation in the value of this displacement of the Bosphorus database is large
382 (0.32 cm) compared to our results (0.02 cm). This is due to this value being estimated from
383 65 different subjects, while our result was only estimated using one patient-specific model
384 of the face.



385

386 *Figure 14. Displacement of the corner point of the mouth (moving up direction) of our*
 387 *method and from the Bosphorus database of the smile position compared to the neutral*
 388 *position.*

389 **4. Discussion**

390 Understanding muscle coordination mechanism of facial expressions plays a crucial role in
 391 the facial rehabilitation interventions for facial palsy or facial transplantation patients.
 392 Numerical models (i.e. finite element models) have been intensively developed [31] to
 393 provide a better understanding of this complex process. However, these developed models
 394 are descriptive and their predictive capacity is still limited. Besides, computer-based
 395 monitoring systems, that automatically recognize action units (AUs) to provide
 396 quantitative and objective information on the facial motion during the rehabilitation
 397 exercise have been developed [5, 36, 37]. Despite many efforts, understanding of facial
 398 motion mechanism still remains a scientific and clinical challenge to help the involved
 399 patients to recover functional facial movements. In particular, the role of muscle excitation

400 and its value for performing a desired facial movement for facial rehabilitation is still an
401 open and longstanding research question. To achieve this complex and challenging
402 objective, the present study aimed to couple reinforcement learning approach to a muscle-
403 driven biomechanical model of the face to explore the facial motion learning capacity such
404 as symmetry and facial smiling actions. For the first time, facial expressions (e.g. smile)
405 are simulated without a priori input data (e.g. motion capture data). In fact, our novel
406 coupling scheme allows to explore emerging properties of the facial muscle contraction
407 mechanism and guides iteratively the face to express smiling action or becomes a
408 symmetric face. Thus, obtained outcomes showed the potential application of this novel
409 approach with facial palsy patients for a better understanding of facial muscle coordination
410 and muscle activation patterns to target a specific motion.

411 More precisely, regarding the symmetric motion of the face, the muscle contraction
412 involves right OOM, right RIS, and right ZYG, while only the left OOP muscle is activated.
413 This is reasonable due to the physical-based model of the face used for symmetry training
414 is drooping of the mouth on the right side. From the biomechanics point of view, this is a
415 symptom of the facial palsy patient on the affected side of the face. In smile-oriented
416 motion of the face, levator anguli oris and zygomaticus are the main muscles responsible
417 for smile action resulting in two corner points in the mouth moving up 0.35 cm, which is
418 within the range of motion compared to the Bosphorus database (0.4 ± 0.32 cm). There is
419 also a good agreement in muscle involvement for smile training as LAO and ZYG
420 compared with the simulation of Flynn et al. [16]. Note that Flynn et al. manually adjusted
421 muscle excitation value to find the appropriate value for expression movements of the
422 finite element model of the face. In fact, our present study revealed the usefulness of the
423 mechanical modeling coupled with reinforcement learning to guide the design of patient

424 specific precision rehabilitation for the face with muscle activation and coordination
425 mechanisms.

426 Recently, deep reinforcement learning becomes an interesting solution for complex control
427 problems [33]. The coupling between a reinforcement learning strategy and a deep neural
428 network allows the agent to build knowledge by gathering information while interacting
429 with the environment. In fact, no prior data is required for training. This particular
430 character enhances the predictive capacity of the involved model. Indeed, reinforcement
431 learning methods were used to solve our problem since they require no prior input data (i.e.
432 muscle excitation and activation patterns). This is particularly useful for learning muscle-
433 driven facial motion problems. While other learning approaches require a database for
434 training, collecting experimental data by directly measuring muscle excitations from living
435 subjects is hard or impossible due to safety and accessibility limitations [9]. One of the
436 challenges when developing an efficient RL model relates to the use cumulative rewards to
437 quantify how the agents ought to take actions in an environment. In our present study,
438 specific rewards were defined to guide the motion patterns toward the specific targets
439 (symmetry and smiling motions). The efficiency of training reinforcement learning
440 algorithms is demonstrated by the reward values predicted using the learned model. The
441 "proper" reward function must be defined for successful training reinforcement learning.
442 The current work used two simple reward functions to train the finite element model of the
443 face for symmetry and smile motion. For more realistic outcomes, a bioinspired reward
444 function built from the bio-mechanical knowledge should be developed for reshaping the
445 mechanism of the desired facial motion. Moreover, two the state-of-the-art RL methods
446 (DDPG method and the successor TD3), which are off-policy algorithms and applicable
447 for complex environments with continuous action spaces, were used to drive the face

448 toward the targeted motions from the activation of the facial muscles. Note that the use of
449 these methods leads to the win of the Learn to Move competition [35]. The reward values
450 during training model using DDPG algorithm seem noisier compared to that of training
451 model using TD3 algorithm. This is due to the nature of DDPG algorithm when it updating
452 the policy more often during training. In terms of reward predicted from the trained model,
453 TD3 seems to have slightly better performance compared to DDPG.

454 One of the most important limitations of the present work deals with the sensitive nature of
455 the hyperparameters of the physics-based face model. Further investigations should be
456 done to take the uncertainties of these parameters into account to provide more reliable
457 prediction outcomes. In fact, each parameter should be represented in a more generic
458 format like interval or probability-based structures (e.g. probability density function (PDF);
459 cumulative distribution function (CDF)) and then associated outcomes (i.e. muscle
460 activation) should be estimated within a plausible range of values. However, taking the
461 parameter uncertainty into account increases drastically the computational cost during the
462 reinforcement learning process. Thus, more efficient uncertainty propagation algorithms
463 should be investigated to scope with this constraint. Moreover, the present face model
464 includes only 10 muscles. In particular, all parameters were set up for a generic model.
465 Thus, a more detailed face model and patient-specific properties of the face tissues and
466 structures should be taken into consideration from medical imaging toward a patient
467 specific rehabilitation application. In particular, the increase of the number of muscles of
468 interest will allow the modeling system to explore full muscle action patterns of the face.
469 Thus, the present system could benefit from the FACS pattern for a given expression to
470 converge quickly to the optimal solution and then other applications like speech synthesis
471 or language learning could be investigated. Regarding the limitation of the used RL

472 approach, the use of only a deep neural network seems to be underestimated for the
473 complex face motion coordination. As perspective, a multi-network approach should be
474 investigated for a better coordination of the facial muscle activations and contractions.
475 Finally, the coupling between RL and FE modeling frameworks requires the development
476 of a specific communication protocol. In a further work, the developed information
477 exchange protocol will be improved to provide a generic communication channel between
478 the RL framework and any other powerful and dedicated FE modeling frameworks like
479 Abaqus or Ansys to overcome the limitation of the current physics-based face model.

480 **5. Conclusions**

481 The present study explored the muscle excitation patterns by coupling reinforcement
482 learning with a finite element model of the face. We developed, for the first time, a novel
483 coupling scheme to integrate the finite element simulation into the reinforcement learning
484 for facial motion learning. In particular, two state-of-the-art reinforcement learning
485 algorithms (deep deterministic policy gradient (DDPG) and Twin-delayed DDPG (TD3))
486 were successfully applied and implemented to drive the simulations of symmetry-oriented
487 and smile movements. Obtained results were in very good agreement with experimental
488 observation. In fact, a better understanding of the facial muscle activation and coordination
489 mechanism is of great clinical interest to guide the optimal rehabilitation strategy. The
490 present work opens new avenues to achieve this challenging objective. As perspectives,
491 this present workflow will be applied for facial palsy and facial transplantation patients to
492 guide and optimize the functional rehabilitation program.

493 **Acknowledgement**

494 This work was financially supported by Sorbonne Center for Artificial Intelligence (SCAI).

495 **Compliance with ethical standards**

496 The authors declare that there are no ethical issues for the present work.

497 **References**

- 498 [1] Bogart, Kathleen R., Linda Tickle-Degnen, and Nalini Ambady. "Communicating
499 without the face: holistic perception of emotions of people with facial paralysis." *Basic and*
500 *applied social psychology* 36.4 (2014): 309-320.
- 501 [2] Magagna, Jeanne. "Communicating without words." *The Silent Child*. Routledge, 2018.
502 29-46.
- 503 [3] Fuller, Geraint, and Cathy Morgan. "Bell's palsy syndrome: mimics and
504 chameleons." *Practical Neurology* 16.6 (2016): 439-444.
- 505 [4] Grewal, D. S. "Atlas of Surgery of the Facial Nerve: An Otolaryngologist's
506 Perspective". *JAYPEE BROTHERS PUBLISHERS*, 2014.
- 507 [5] Haase, Daniel, et al. "Automated and objective action coding of facial expressions in
508 patients with acute facial palsy." *European Archives of Oto-Rhino-Laryngology* 272.5
509 (2015): 1259-1267.
- 510 [6] He, Shu, et al. "Quantitative analysis of facial paralysis using local binary patterns in
511 biomedical videos." *IEEE Transactions on Biomedical Engineering* 56.7 (2009): 1864-
512 1870.
- 513 [7] Robinson, Mara Wernick, et al. "Facial rehabilitation." *Operative Techniques in*
514 *Otolaryngology-Head and Neck Surgery* 23.4 (2012): 288-296.
- 515 [8] Jayatilake, Dushyantha, et al. "Robot assisted physiotherapy to support rehabilitation of
516 facial paralysis." *IEEE Transactions on Neural Systems and Rehabilitation*
517 *Engineering* 22.3 (2013): 644-653.

- 518 [9] Pileicikiene, Gaiville, et al. "A three-dimensional model of the human masticatory
519 system, including the mandible, the dentition and the temporomandibular
520 joints." *Stomatologija* 9.1 (2007): 27-32.s
- 521 [10] Erdemir, Ahmet, et al. "Model-based estimation of muscle forces exerted during
522 movements." *Clinical biomechanics* 22.2 (2007): 131-154.
- 523 [11] Zhang, Yu, Edmond C. Prakash, and Eric Sung. "Face alive." *Journal of Visual*
524 *Languages & Computing* 15.2 (2004): 125-160.
- 525 [12] Claes, Peter, et al. "Computerized craniofacial reconstruction: conceptual framework
526 and review." *Forensic science international* 201.1-3 (2010): 138-145.
- 527 [13] Mollemans, Wouter, et al. "Predicting soft tissue deformations for a maxillofacial
528 surgery planning system: from computational strategies to a complete clinical
529 validation." *Medical image analysis* 11.3 (2007): 282-301.
- 530 [14] Kim, Hyungmin, et al. "A new soft-tissue simulation strategy for cranio-maxillofacial
531 surgery using facial muscle template model." *Progress in biophysics and molecular*
532 *biology* 103.2-3 (2010): 284-291.
- 533 [15] Hannam, A. G. "Current computational modelling trends in craniomandibular
534 biomechanics and their clinical implications." *Journal of oral rehabilitation* 38.3 (2011):
535 217-234.
- 536 [16] Flynn, C., Stavness, I., Lloyd, J., & Fels, S. (2015). A finite element model of the face
537 including an orthotropic skin model under in vivo tension. *Computer methods in*
538 *biomechanics and biomedical engineering*, 18(6), 571-582.

- 539 [17] Yu, Chao, Jiming Liu, and Shamim Nemati. "Reinforcement learning in healthcare: A
540 survey." *arXiv preprint arXiv:1908.08796* (2019).
- 541 [18] Kormushev, Petar, Sylvain Calinon, and Darwin G. Caldwell. "Reinforcement
542 learning in robotics: Applications and real-world challenges." *Robotics 2.3* (2013): 122-
543 148.
- 544 [19] Szita, István. "Reinforcement learning in games." *Reinforcement learning*. Springer,
545 Berlin, Heidelberg, 2012. 539-577.
- 546 [20] Sallab, Ahmad EL, et al. "Deep reinforcement learning framework for autonomous
547 driving." *Electronic Imaging 2017.19* (2017): 70-76.
- 548 [21] Bueno, M. B., Nieto, X. G. I., Marqués, F., & Torres, J. (2017). Hierarchical object
549 detection with deep reinforcement learning. *Deep Learning for Image Processing*
550 *Applications*, 31(164), 3.
- 551 [22] Jonsson, A. (2019). Deep reinforcement learning in medicine. *Kidney Diseases*, 5(1),
552 18-22.
- 553 [23] Gottesman, Omer, et al. "Guidelines for reinforcement learning in healthcare." *Nature*
554 *medicine* 25.1 (2019): 16-18.
- 555 [24] Maia, Tiago V., and Michael J. Frank. "From reinforcement learning models to
556 psychiatric and neurological disorders." *Nature neuroscience* 14.2 (2011): 154-162.
- 557 [25] Nazari, Mohammad Ali, et al. "Simulation of dynamic orofacial movements using a
558 constitutive law varying with muscle activation." *Computer methods in biomechanics and*
559 *biomedical engineering* 13.4 (2010): 469-482.

- 560 [26] Bucki, Marek, Mohammad Ali Nazari, and Yohan Payan. "Finite element speaker-
561 specific face model generation for the study of speech production." *Computer methods in*
562 *biomechanics and biomedical engineering* 13.4 (2010): 459-467.
- 563 [27] John E. Lloyd, Ian Stavness, and Sidney Fels, "ArtiSynth: A fast interactive
564 biomechanical modeling toolkit combining multibody and finite element simulation", *Soft*
565 *Tissue Biomechanical Modeling for Computer Assisted Surgery*, pp. 355-394, Springer,
566 2012.
- 567 [28] Stavness, Ian, et al. "Coupled biomechanical modeling of the face, jaw, skull, tongue,
568 and hyoid bone." *3D multiscale physiological human*. Springer, London, 2014. 253-274.
- 569 [29] Jiang, Zeqing, et al. "A Cloud-Based Training and Evaluation System for Facial
570 Paralysis Rehabilitation." *2018 IEEE 16th International Conference on Industrial*
571 *Informatics (INDIN)*. IEEE, 2018.
- 572 [30] Abdi, Amir H., et al. "Reinforcement learning for high-dimensional continuous
573 control in biomechanics: an intro to artisynth-rl." *arXiv preprint arXiv:1910.13859* (2019).
- 574 [31] Fan, A. X., Dakpé, S., Dao, T. T., Pouletaut, P., Rachik, M., & Ho Ba Tho, M. C.
575 (2017). "MRI-based finite element modeling of facial mimics: a case study on the paired
576 zygomaticus major muscles." *Computer methods in biomechanics and biomedical*
577 *engineering*, 20(9), 919-928.
- 578 [32] Dao, T. T., Fan, A. X., Dakpe, S., Pouletaut, P., Rachik, M., & HO BA THO, M. C.
579 (2018). "Image-based skeletal muscle coordination: case study on a subject specific facial
580 mimic simulation." *Journal of Mechanics in Medicine and Biology*, 18(02), 1850020.

- 581 [33] K Nowakowski, P Carvalho, J B Six, Y Maillet, A T Nguyen, I Seghiri, L M'Pemba,
582 T Marcille, S T Ngo, T T Dao (2021). "Human locomotion with reinforcement learning
583 using bioinspired reward reshaping strategies." *Medical & Biological Engineering &*
584 *Computing*, 59(1), 243-256.
- 585 [34] Asri, Hiba, et al. "Big data in healthcare: Challenges and opportunities." *2015*
586 *International Conference on Cloud Technologies and Applications (CloudTech)*. IEEE,
587 2015.
- 588 [35] Song, Seungmoon, et al. "Deep reinforcement learning for modeling human
589 locomotion control in neuromechanical simulation." *Journal of NeuroEngineering and*
590 *Rehabilitation* 18.1 (2021): 1-17. <https://doi.org/10.1186/s12984-021-00919-y>
- 591 [36] Hamm, Jihun, et al. "Automated facial action coding system for dynamic analysis of
592 facial expressions in neuropsychiatric disorders." *Journal of neuroscience methods* 200.2
593 (2011): 237-256.
- 594 [37] Tian, Y-I., Takeo Kanade, and Jeffrey F. Cohn. "Recognizing action units for facial
595 expression analysis." *IEEE Transactions on pattern analysis and machine intelligence* 23.2
596 (2001): 97-115.
- 597 [38] Choi, Kang Young. "Analysis of facial asymmetry." *Archives of craniofacial*
598 *surgery* 16.1 (2015): 1
- 599 [39] Eskes, Merijn, et al. "Simulation of facial expressions using person-specific sEMG
600 signals controlling a biomechanical face model." *International journal of computer*
601 *assisted radiology and surgery* 13.1 (2018): 47-59.

- 602 [40] Nguyen, Tan-Nhu, et al. "Kinect-driven patient-specific head, skull, and muscle
603 network modelling for facial palsy patients." *Computer Methods and Programs in*
604 *Biomedicine* 200 (2021): 105846.
- 605 [41] Fujimoto, Scott, Herke Hoof, and David Meger. "Addressing function approximation
606 error in actor-critic methods." *International conference on machine learning*. PMLR, 2018.

607 **AUTHOR BIOGRAPHIES**



Duc-Phong NGUYEN is currently a Ph.D student in Biomechanic Engineering at the Uninversity of Technology of Compiègne (France). His current research interest is facial recognition and rehabilitation for medical computer aided system diagnosis.



Marie-Christine HO BA THO is Full Professor in Mechanics since 1998 at the Uninversity of Technology of Compiègne (France). She is currently Head of Research at UTC. Her research work concerns Biomechanics of the musculoskeletal system and the modelling of uncertainties of biomechanics data.



Tien Tuan Dao is Full Professor of Biomedical Engineering and Biomechanics at Centrale Lille Institut, France. His research interests concern computational biomechanics, knowledge and system engineering, and in silico medicine.

