

Restricted Boltzmann Machines and Their Extensions for Face Modeling

Chi Nhan Duong and Khoa Luu*

Carnegie Mellon University, USA

Received: August 22, 2017; Published: September 08, 2017

*Corresponding author: Khoa Luu, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA, Email: chinhand@andrew.cmu.edu

Abstract

Restricted Boltzmann Machines, Deep Boltzmann Machines, and their extensions have brought much attention and become powerful tools for many machine learning tasks. The increased popularity of these techniques is not only limited in modelling static data but also time-series data. In this paper, we aim to give a review of recent developments of such models for sequential data modelling. Their structures, energy functions, learning algorithm as well as applications are also provided with systematic discussions.

Keywords: Boltzmann Machines; Face Modeling; Longitudinal Face Modeling

Abbreviations: TRBM: Temporal Restricted Boltzmann Machines; RBM: Restricted Boltzmann Machines; SGD: Stochastic Gradient Descent; BB-RBM: Binary-Binary RBM; CRM: Conditional RBM; GCRBM: Gated Conditional Restricted Boltzmann Machines; TRBM: Temporal Restricted Boltzmann Machine; DBM: Deep Boltzmann Machines; AGFW: Aging Faces in the Wild; CACD: Cross-Age Celebrity Dataset

Introduction

In this paper, we review the main structure of Restricted Boltzmann Machines and their improvements that have been proposed for data modeling. Different applications of this model are also summarized to provide a better understanding of the model. Notice that our work is different from the survey of recent developments of feature learning algorithms for time-series problems. The aim of that survey is to review unsupervised feature learning methods together with applications to several classical time-series problems. On the other hand, in this work, we also review recent trends and extensions of a specific model, i.e. the Temporal Restricted Boltzmann Machines (TRBM), for face modeling.

Restricted Boltzmann Machines

Restricted Boltzmann Machines (RBM) [1,2] are undirected graphical models consisting two layers of stochastic units, i.e. visible v and hidden units h . This is a simplified version of Boltzmann Machines where no intra connection between units in the same layer is created. RBM structure is a bipartite graph where visible and hidden units are pair wise conditionally independent. Given a binary state of $\{v, h\}$, the energy of RBM and the joint distribution of visible and hidden units can be computed as.

$$-E(v, h) = v^T W h + b^T v + a^T h$$

$$P(v, h; \theta) = \frac{1}{Z(\theta)} \exp\{-E(v, h)\} \quad (1)$$

Where $\theta = \{W, b, a\}$ denotes the parameter set of RBM including the connection weights; and the biases of visible and hidden units, respectively. In contrast to general BM, the inference process in RBM is exact. Thanks to this important property, the Contrastive Divergence technique can perform well to obtain the model parameters for RBM.

The conditional probabilities RBM structure can be computed as follows.

$$p(h_j = 1 | v) = \sigma\left(\sum_i v_i w_{ij} + a_j\right)$$

$$p(v_i = 1 | h) = \sigma\left(\sum_j v_j w_{ji} + a_i\right) \quad (2)$$

Where $\sigma(\cdot)$ is the logistic function.

In the training process of RBM, its parameters are learned to minimize the log-likelihood $\theta^* = \arg \max \log P(v, h; \theta)$. The Stochastic Gradient Descent (SGD) procedure is usually employed for this optimization. The gradients are given by

$$\frac{\partial}{\partial \theta} \mathbb{E}[\log P(v, h; \theta)] = \mathbb{E}_{data} \left[\frac{\partial E}{\partial \theta} \right] - \mathbb{E}_{model} \left[\frac{\partial E}{\partial \theta} \right] \quad (3)$$

Where $\mathbb{E}_{data} \left[\frac{\partial E}{\partial \theta} \right]$ and $\mathbb{E}_{model} \left[\frac{\partial E}{\partial \theta} \right]$ are the expectations with respect to data distribution and distribution estimated by the RBM.

Extensions of Restricted Boltzmann Machines

In the original RBM, both visible and hidden units are binary. Therefore, it is also called as Binary-Binary RBM (BB-RBM). To

make it more powerful and be able to deal with different kinds of data, RBM has received several extensions in its structures and unit

types (i.e. binary, linear). This section will revise the main features of these extensions (Figure 1).

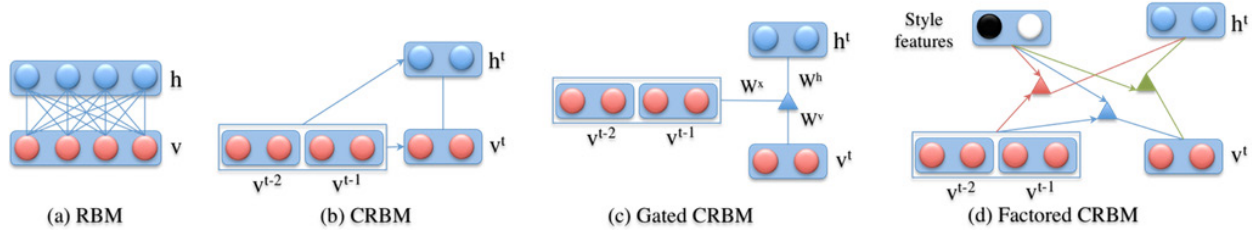


Figure 1: Restricted Boltzmann Machines and its variants.

Gaussian Restricted Boltzmann Machines: Instead of using the binary visible units as the original RBM, Gaussian RBM [3] assumes the visible units have values in $[-\infty, \infty]$ and normally distributed with mean b_i and variance σ_i^2 . By this way, this extension of RBM can be used for modelling real-valued data, i.e. pixel intensities. The energy function is modified as follows.

$$E(v, h; \theta) = \sum_i \frac{(v_i - b_i)^2}{2\sigma_i^2} - \sum_{ij} \frac{v_i}{\sigma_i} w_{ij} h_j - \sum_j a_j h_j \quad (4)$$

The conditional distributions over “v” and h are then given as in Eqn. (5).

$$p(h_j | v) = \delta \left(\sum_i w_{ij} \frac{v_i}{\sigma_i} + a_j \right)$$

$$v_i | h \sim N(\sigma_i \sum_j w_{ij} h_j + b_i, \sigma_i^2) \quad (5)$$

Conditional Restricted Boltzmann Machines

Focusing more on the RBM structure, Taylor et al. [4] introduced the Conditional RBM (CRM) to exploit the temporal relationship between consecutive frames in time-series data. Two more interactions for RBM structure, i.e. visible-to-visible and visible-to-hidden feed forward connections from previous time steps are added to the standard RBM structure. These connections are effectively incorporated to the formulation of bias terms making the learning process similar to RBM. Moreover, it can inherit most important properties of standard RBM such as simple, exact inference and efficient approximate learning. Later, Hinton [5] refined the CRBM structure by implementing the multiplicative interactions in their proposed Gated Conditional Restricted Boltzmann Machines (GCRBM). In this model, three sets of units, i.e. input, output and hidden units are first defined. Then instead of simply incorporating them via bias terms as CRBM, a three-way interaction is introduced to let the input units directly influence the interactions between units. By this way, the input units will be able to gate the basic function for reconstructing the output. The energy function is defined as

$$-E(v, h | x, \theta) = \sum_{ijk} w_{ijk} v_i h_j x_k + \sum_{ij} w_{ij} v_i h_j + \sum_i b_i v_i + \sum_j a_j h_j \quad (6)$$

With a set of deterministic factors f, w_{ijk} can be factorized to three pair wise interactions, (1) w_{ij}^v connects v_i to factor f; (2) w_{jf}^h connects h_j to factor f; (3) w_{kf}^x connects x_k to factor f.

There are several other extensions following the structure of CRBM such as factored CRBM [6] for motion style modeling; [7] with additional hierarchical structure for style interpolation; implicit Mixtures of CRBM (imCRBM) [8] that can learn from the data with several modes (e.g. walking and running in body modeling task).

Temporal Restricted Boltzmann Machines

Temporal Restricted Boltzmann Machine (TRBM) [9] has gained significant attention as one of the probabilistic model that can accurately model complex time-series structure while keeping the inference tractable. The major difference between the original RBM and TRBM is the directed connections from both visible and hidden units of previous states to the current states. With these new connections, the short history of their activations can act as “memory” and is able to contribute to the inference step of current states of visible units. The joint distribution over (v^t, h^t) at time t is conditional on the past m states and given as

$$p(v^t, h^t | v_{t-m}^{t-1}) = \frac{\exp(-E(v^t, h^t | v_{t-m}^{t-1}, h_{t-m}^{t-1}))}{z(v_{t-m}^{t-1}, h_{t-m}^{t-1})} \quad (7)$$

where $(v_{(t-m)}^{(t-1)}, h_{(t-m)}^{(t-1)})$ are the sequence of visible and hidden units from time t-m to t-1. The energy is computed as in Eqn. (1) except the new bias terms are defined as

$$\hat{b}^t = b + \sum_{k=1}^m B_k v^{t-k}$$

$$\hat{a}^t = a + \sum_{k=1}^m A_k h^{t-k} + \sum_{k=1}^m c_k v^{t-k} \quad (8)$$

Sutskever et al. [10] later proposed the Recurrent TRBM (RTRBM) to overcome the issues of non-trivial inference step in TRBM. In RTRBM, instead of using binary hidden units, a real variable $r^t = E[h^t | v^t]$ is introduced and incorporated to the energy function:

$$-E(v^t, h^t | r^{t-1}) = h^{tT} W v^t + \hat{b}^T v^t + a^T h^t + h^{tT} W r^{t-1} \quad (9)$$

Where denotes the weights of the connections from r^{t-1} to h^t and $r^t; r^t = \sigma(W v^t + \hat{a} + W r^{t-1})$ if $t > 1$ and $r^t = \sigma(W v^t + \hat{a}_{init})$ if $t = 1$.

Other TRBM extensions consists of employing denoising Auto encoder to initialize the weights for the hidden-to-hidden connections [11]; transferring information between two sequences

with Input-Output TRBM (IOTRBM) and Factor Third-order Input-Output TRBM (FIOTRBM) [12]; or removing the unnecessary connections from the fully connected topology between visible and hidden units of the RTRBM. Structured RTRBM (SRTRBM) [13]. Furthermore, in SRTRBM, the main idea is to they constructed block masking M_w and M_w for the weight matrices W and W

to model sparsely connectivity between groups of visible units and hidden units; and between groups of hidden units themselves. The authors also suggested to use spike and slab RBM instead of the Gaussian RBM for better conditional covariance modeling (Figure 2).

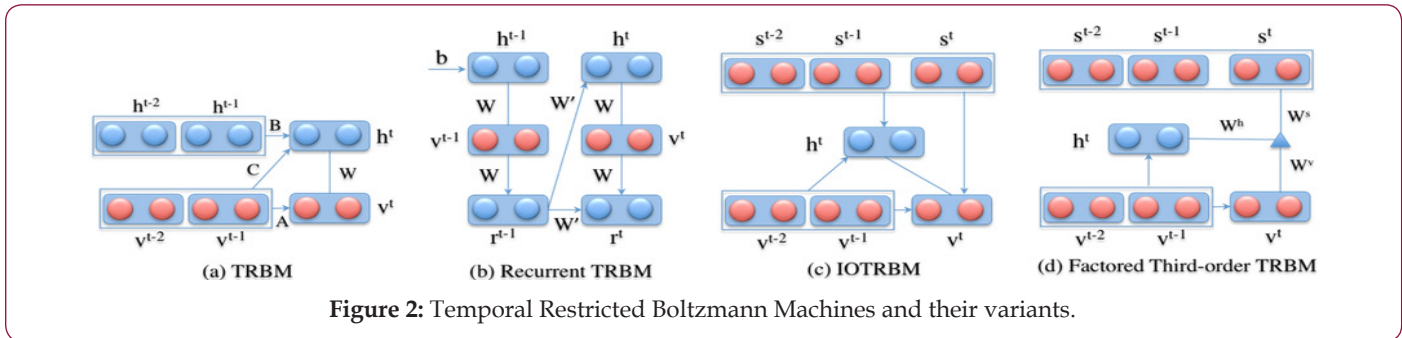


Figure 2: Temporal Restricted Boltzmann Machines and their variants.

Deep Boltzmann Machines for Face Modeling

Deep Boltzmann Machines (DBM) [14] is a probabilistic generative model that consists of many hidden layers. Each higher layer plays a role of capturing the correlations between features of its lower layer. The structure of DBM contains several RBMs are organized in a layered manner. In DBM, the connections are between visible units and the hidden units in the first layer as well as between the hidden units in adjacent hidden layers. Thanks to this structure, the hidden units in higher layer can learned more complicated correlations of features captured in lower layer. Another interesting point of DBM is that these higher representations can be built from the training data in an unsupervised fashion. Then the labeled training data, which is assumed to be very limited, can be used to fine tune the model for a particular application. Notice that unlike other models such as Deep Belief Network [15] or Deep Auto encoders [16], all connections between units in two consecutive layers are undirected. As a result, each unit receives both bottom-up and top-down information and, therefore, better propagate uncertainty during the inference process.

Let "v" be the set of visible units and $\{h^{(1)}, h^{(2)}, h^{(3)}\}$ be the set of units in three hidden layers, the energy of the state $\{v, h^{(1)}, h^{(2)}, h^{(3)}\}$ is given as follows.

$$-E(v, h^{(1)}, h^{(2)}, h^{(3)}; \theta) = v^T w^{(1)} h^{(1)} + h^{(1)T} w^{(2)} h^{(2)} + h^{(2)T} w^{(3)} h^{(3)} \quad (10)$$

Where $\theta = \{W^{(1)}, W^{(2)}, W^{(3)}\}$ are the weights of visible-to-hidden and hidden-to-hidden connections. Notice that the bias terms for visible and hidden units are ignored in Eqn. (10) for simplifying the representation. Similar to RBM, the probability of a visible vector "v" assigned by the model is

$$p(v; \theta) = \frac{1}{Z(\theta)} \sum_{h^{(1)}, h^{(2)}, h^{(3)}} \exp(-E(v, h^{(1)}, h^{(2)}, h^{(3)}; \theta)) \quad (11)$$

DBM is shown to be more robust with ambiguous input data [14]. Besides some recent works using DBM as shape prior model [8,17,18]. Duong et al. [19] proposed the Deep Appearance Models (DAM) for both tasks of learning high-level representation and face reconstruction under various challenging conditions. In DAM

structure, both shape and texture are modeled using two different DBMs. Then on the top of these two deep models, the higher-level relationships of both shape and texture are exploited so that the reconstruction of one can benefit from the information on the other. Three crucial components represented in hierarchical layers are modeled to robustly capture the variations of facial shapes and appearances. DAM is therefore superior to other classical models such as Active Appearance Models (AAM) in differencing a representation for new face images under various conditions. A robust version of DAM, named Robust Appearance Models (RDAM), is later introduced in [20] to extend its ability of dealing with occluded face regions. In the structure of RDAM, additional appearance masks is learned and help the DAM to separate corrupted/ occluded pixels in texture modeling process. As a result, those regions are ignored during face reconstruction and model fitting and, therefore, better reconstructed results can be achieved.

In addition to single face modeling, a TRBM based age progression model is introduced in [21] to embed the temporal relationship between images in a face sequence. By taking the advantages of log-likelihood objective function and avoiding the ℓ_2 reconstruction error during training, the proposed model is able to efficiently capture the non-linear aging process and automatically synthesize a series of age-progressed faces in various age ranges with more aging details. In this approach, the long-term age progresses is first decomposed into a sequence of short-term changes and model it as a face sequence. The Temporal Restricted Boltzmann Machines based age progression model together with the prototype faces are then constructed to learn the aging transformation between faces in the sequence. In addition, to enhance the wrinkles of faces in the later age ranges, the wrinkle models are further constructed using Restricted Boltzmann Machines to capture their variations in different facial regions. The geometry constraints are also taken into account in the last step for more consistent age-progressed results.

Recently, addressing a limitation of intractable learning process of TRBM based model, the Temporal Non-Volume Preserving

(TNVP) approach is presented in [22] for embedding the feature transformations between faces in consecutive stages while keeping a tractable density function, exact inference and evaluation. This model shows its advantages not only in capturing the non-linear age-related variance in each stage but also producing a smooth synthesis in age progression across faces. Moreover, the structure can be transformed into a deep convolution network while keeping the advantages of probabilistic models with tractable log-likelihood density estimation. This approach is evaluated in both terms of synthesizing age-progressed faces and cross-age face verification and consistently shows the state-of-the-art results in various face aging databases, i.e. FG-NET, MORPH, our collected large-scale aging database named Aging Faces in the Wild (AGFW), and Cross-Age Celebrity Dataset (CACD).

References

- Hench LL (2006) The story of Bioglass. *J Mater Sci Mater Med* 17(11): 967-978.
- Hench LL, Splinter RJ, Allen WC, Greenlee TK (1971) Bonding mechanisms at the interface of ceramic prosthetic materials. *J Biomed Mater Res* 5(6): 117-141.
- Hench LL, Paschall HA (1973) Direct chemical bond of bioactive glass-ceramic materials to bone and muscle. *J Biomed Mater Res* 7(3): 25-42.
- Baino F, Novajra G, Miguez-Pacheco V, Boccaccini AR, Vitale-Brovarone C (2016) Bioactive glasses: Special applications outside the skeletal system. *J Non-Cryst Solids* 432: 15-30.
- Jones JR (2013) Review of bioactive glass: From Hench to hybrids. *Acta Biomater* 9(1): 4457-4486.
- Jones JR, Brauer DS, Hupa L, Greenspan DC (2016) Bioglass and Bioactive Glasses and Their Impact on Healthcare. *Int J Applied Glass Sci* 7(4): 423-434.
- Sepulveda P, Jones JR, Hench LL (2002) *In vitro* dissolution of melt-derived 45S5 and sol-gel derived 58S bioactive glasses. *J Biomed Mater Res* 61(2): 301-311.
- Vogel M, Voigt C, Gross UM, Muller-Mai CM (2001) *In vivo* comparison of bioactive glass particles in rabbits. *Biomaterials* 22(4): 357-362.
- Clupper DC, Hench LL (2003) Crystallization kinetics of tape cast bioactive glass 45S5. *J Non-Cryst Solids* 318(1): 43-48.
- Chen QZ, Thompson ID, Boccaccini AR (2006) 45S5 Bioglass[®]-derived glass-ceramic scaffolds for bone tissue engineering. *Biomaterials* 27(11): 2414-2425.
- Boccaccini AR, Chen Q, Lefebvre L, Gremillard L, Chevalier J (2007) Sintering, crystallization and biodegradation behaviour of Bioglass[®]-derived glass-ceramics. *Faraday Discussions* 136: 27-44.
- Lefebvre L, Chevalier J, Gremillard L, Zenati R, Thollet G, et al. (2007) Structural transformations of bioactive glass 45S5 with thermal treatments. *Acta Materialia* 55: 3305-3313.
- Lefebvre L, Gremillard L, Chevalier J, Zenati R, Bernache-Assolant D (2008) Sintering behaviour of 45S5 bioactive glass. *Acta Biomater* 4(6): 1894-1903.
- Huang R, Pan J, Boccaccini AR, Chen QZ (2008) A two-scale model for simultaneous sintering and crystallization of glass-ceramic scaffolds for tissue engineering. *Acta Biomater* 4(4): 1031095-1031103.
- Bretcanu O, Chatzistavrou X, Paraskevopoulos K, Conradt R, Thompson I, et al. (2009) Sintering and crystallisation of 45S5 Bioglass[®] powder. *J Eur Ceram Soc* 29(16): 3299-3306.
- Wallace KE, Hill RG, Pembroke JT, Brown CJ, Hatton PV (1999) Influence of sodium oxide content on bioactive glass properties. *J Mater Sci Mater Med* 10(12): 697-701.
- Kansal I, Reddy A, Muñoz F, Choi S-J, Kim H-W, et al. (2014) Structure, biodegradation behavior and cytotoxicity of alkali-containing alkaline-earth phosphosilicate glasses. *Mater Sci Eng C Mater Biol Appl* 44: 159-165.
- Brink M, Turunen T, Happonen RP, Yli-Urpo A (1997) Compositional dependence of bioactivity of glasses in the system Na₂O-K₂O-MgO-CaO-B₂O₃-P₂O₅-SiO₂. *J Biomed Mater Res* 37(1): 114-121.
- Lusvardi G, Malavasi G, Menabue L, Aina V, Morterra C (2009) Fluoride-containing bioactive glasses: surface reactivity in simulated body fluids solutions. *Acta Biomater* 5(9): 3548-3562.
- Lusvardi G, Zaffe D, Menabue L, Bertoldi C, Malavasi G, Consolo U (2008) *In vitro* and *in vivo* behaviour of zinc-doped phosphosilicate glasses. *Acta Biomater* 5(1): 419-428.
- Zhang D, Hupa M, Hupa L (2008) In situ pH within particle beds of bioactive glasses. *Acta Biomater* 4(5): 1498-1505.
- Waltimo T, Brunner TJ, Vollenweider M, Stark WJ, Zehnder M (2007) Antimicrobial effect of nanometric bioactive glass 45S5. *J Dental Res* 86(8): 754-757.
- Chen Q-Z, Li Y, Jin L-Y, Quinn JMW, Komesaroff PA (2010) A new sol-gel process for producing Na₂O-containing bioactive glass ceramics. *Acta Biomater* 6(10): 4143-4153.
- Brauer DS, Karpukhina N, O Donnell MD, Law RV, Hill RG (2010) Fluoride-containing bioactive glasses: Effect of glass design and structure on degradation, pH and apatite formation in simulated body fluid. *Acta Biomater* 6(8): 3275-3282.
- Ferreira JMF, Goel A (2016) Alkali-free bioactive glass composition. Google PatentsUS 9238044 B2.
- Goel A, Kapoor S, Rajagopal RR, Pascual MJ, Kim H-W, et al. (2012) Alkali-free bioactive glasses for bone tissue engineering: a preliminary investigation. *Acta Biomater* 8(1): 361-372.
- Kansal I, Goel A, Tulyaganov DU, Pascual MJ, Lee HY, et al. (2011) Diopside-fluorapatite glass-ceramics: Potential materials for bone tissue engineering. *J Mater Chem* 21(40): 16247-16256.
- Kansal I, Goel A, Tulyaganov DU, Rajagopal RR, Ferreira JMF (2012) Structural and thermal characterization of CaO-MgO-SiO₂-P₂O₅-CaF₂ glasses. *J Eur Cer Soc* 32(11): 2739-2746.
- Kapoor S, Semitela Â, Goel A, Xiang Y, Du J, et al. (2015) Understanding the composition-structure-bioactivity relationships in diopside (CaO.MgO.2SiO₂)-tricalcium phosphate (3CaO.P₂O₅) glass system. *Acta Biomater* 15: 210-226.
- Goel A, Kapoor S, Tilocca A, Rajagopal RR, Ferreira JMF (2013) Structural role of zinc in biodegradation of alkali-free bioactive glasses. *J Mater Chem B* 1(24): 3073-3082.
- Goel A, Rajagopal RR, Ferreira JMF (2011) Influence of strontium on structure, sintering and biodegradation behavior of CaO-MgO-SrO-SiO₂-P₂O₅-CaF₂ glasses. *Acta Biomater* 7(11): 4071-4080.
- Kapoor S, Goel A, Tilocca A, Dhuna V, Bhatia G, et al. (2014) Role of glass structure in defining the chemical dissolution behavior, bioactivity and antioxidant properties of zinc and strontium co-doped alkali-free phosphosilicate glasses. *Acta Biomater* 10(7): 3264-3278.
- Kapoor S, Goel A, Pascual MJ, Ferreira JMF (2013) Thermo-mechanical behaviour of alkali free bioactive glass-ceramics co-doped with strontium and zinc. *J Non-Crystal Solids* 375: 74-82.
- Brito AF, Antunes B, dos Santos F, Fernandes HR, Ferreira JMF (2016) Osteogenic capacity of alkali-free bioactive glasses. *In vitro* studies. *J Biomed Mater Res Part B Appl Biomater*, p. 1-6.
- Cortez PP, Brito AF, Kapoor S, Correia AF, Atayde LM, et al. (2017) The *in vivo* performance of an alkali-free bioactive glass for bone grafting, FastOs[®]BG, assessed with an ovine model. *J Biomed Mater Res Part B Appl Biomater* 105(1): 30-38.

36. Susana M Olhero, Hugo R Fernandes, Catarina F Marques, Bianca CG Silva, José MF Ferreira (2017) Additive manufacturing of 3D porous alkali-free bioactive glass scaffolds for healthcare applications. *J Mat Sci* 52(20): 12079-12088.
37. Eqtesadi S, Motealleh A, Miranda P, Lemos A, Rebelo A (2013) A simple recipe for direct writing complex 45S5 Bioglass 3D scaffolds. *Mater Lett* 93: 68-71.
38. Eqtesadi S, Motealleh A, Miranda P, Pajares A, Lemos A, et al. (2014) Robocasting of 45S5 bioactive glass scaffolds for bone tissue engineering. *J Eur Ceram Soc* 34(1): 107-118.



Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

<http://biomedres.us/>