The universal approximation capabilities of 2pi-periodic approximate identity neural networks

Saeed Panahian Fard
School of Mathematical Sciences
Universiti Sains Malaysia
11800 USM Penang, Malaysia
Email: saeedpanahian@yahoo.com

Zarita Zainuddin
School of Mathematical Sciences
Universiti Sains Malaysia
11800 USM Penang, Malaysia
Email: zarita@cs.usm.my

***Abstract*—A fundamental theoretical aspect of artificial neural networks is related to the investigation of the universal approximation capability of a new type of a three-layer feedforward neural networks. In this study, we present four theorems concerning the universal approximation capabilities of a three-layer feedforward 2pi-periodic approximate identity neural networks. Using 2pi-periodic approximate identity, we prove two theorems which show the universal approximation capability of a three-layer feedforward 2pi-periodic approximate identity neural networks in the space of continuous 2pi-periodic functions. The proofs of these theorems are based on the convolution linear operators and the theory of ϵ-net. Using 2pi-periodic approximate identity again, we also prove another two theorems which show the universal approximation capability of these networks in the space of pth-order Lebesgue integrable 2pi-periodic functions.**

***Index Terms*—Universal approximation, 2pi-periodic approximate identity, 2pi-periodic approximate identity neural networks, continuous 2pi-periodic functions, pth-order Lebesgue integrable 2pi-periodic functions, Generalized Minkowski inequality.**

I. Introduction

It is well known that artificial neural networks are very useful tools for approximationg functions [1]. In the literature, the universal approximation capability of various feedforward neural networks (FNNs) has been proved by many authors. Some important research are as follows: In 1989, Cybenko proved the universal approximation capability of feedforward sigmoid neural networks [2]. In 1991, Park and Sandberg proved the universal approximation capability of radial basis function FNNs [3]. In 1998, Scarselli and Tosi reviewed the universal approximation capability of FNNs [4]. Recently, In 2008, Sanguineti surveyed the above property of FNNs [5]. More recently, In 2012, Ting-fan and Xin-long constructed this property of FNNs in the space of continuous function defined on any compact set of multi dimensionals of real numbers [6].

The universal approximation capability of feedforward neural networks (FNNs) deals with the possibility of approximation functions, functionals and operators by these networks. In other words, the universal approximation capability of feedforward neural networks shows that feedforward neural networks can approximate any function, functional and operator to any arbitrary accuracy if a sufficient number of activation functions is available in the networks [7]. In mathematical terms, this property means that these networks are dense in the space of functions, functionals and operators [8].

On the other hand, there have been considerable interest in periodic neural networks in the last few years. In 2010, Hahm and Hong proved the approximation capability of a class of a three-layer feedforward periodic neural networks based on constructive approach [9]. In 2011, Gecynalda et al. made a comparison based on financial market times series between sincos neural networks and sinc neural networks [10]. In 2012, Wang et al. obtained further accuracy of approximation function by periodic neural networks can be higher than previous works [11]. In the same year, Chen et al. constructed periodic sigmoid neural networks [12].

It is well known that approximate identity neural networks are universal approximators [13]. In 2012, we presented the universal approximation capability of double approximate identity neural networks in real Lebesgue spaces [14]. In 2013, we proved that a three layer feedforward flexible approximate identity neural networks are universal approximators in the space of continuous functions and in the space of real Lebesgue functions [15], [16]. Moreover, we also proved the universal approximation capabilities of Mellin approximate identity neural networks [17]. Furthermore, the universal approximation capability of double flexible approximate identity neural networks is shown by these authors [18].

In this study, we are motivated to reduce the gap among the above previous findings. In fact, we solve the following problems: how can the universal approximation capability of a three-layer feedforward 2pi-periodic approximate identity neural networks arise in the space of continuous 2pi-periodic functions? In addition, how can the universal approximation capability of these networks occur in the space of pth-order Lebesgue integrable 2pi-periodic functions?

In order to solve the above problems, we will prove some theorems to show the universal approximation capabilities of a three layer feedforward 2pi-periodic approximate identity neural networks. First, we will recall the definition of the 2pi-periodic approximate identity functions. Second, we will prove a theorem based on the 2pi-periodic approximate identity and the convolution linear operators in the space of continuous 2pi-periodic functions. Third, by using this theorem, we will obtain a main theorem which shows the universal approximation capability of a three-layer 2pi-periodic approximate identity neural networks on any compact subset of the space of

978-1-4799-5245-8/14 $31.00 © 2014 IEEE
DOI 10.1109/ISCC-C.2013.147

continuous 2pi-periodic functions. The proof of this result is in the frame work of theory of ϵ-net and similar to the proof of Wu et al. 's Theorem 1 [19]. Then, we will extend our theoretical results to the space of pth-order Lebesgue integrable 2pi-periodic functions.

This study is organized as follows: in Section II, we review some basic definitions which will be needed in the next sections. In Section III, we present some results concerning the universal approximation capability of a three-layer feedforward 2pi-periodic approximate identity neural networks in the space of continuous 2pi-periodic functions. In Section IV, we give some results related to the universal approximation capability of a three-layer feedforward 2pi-periodic approximate identity neural networks in space of pth-order Lebesgue integrable 2pi-periodic functions. Finally, conclusions are given in Section V.

II. Basic definitions

In this section, we review some basic definitions which will be used in the succeeding sections. As the basic tools of this study, we recall the definition of the 2pi-periodic approximate identity in the following:

***Definition** 2.1:* [20] Let $\{\phi_n(\theta)\}_{n=1}^{\infty}$, $\phi_n : [-\pi, \pi] \to \mathbb{R}$, be a sequence of periodic functions. The sequence is called a 2pi-periodic approximate identity if it satisfies the following conditions:
1) $\frac{1}{2\pi}\int_{-\pi}^{\pi}\phi_n(\theta)d\theta = 1$;
2) for any $\epsilon > 0$ and $0 < \delta < \pi$, there exists a number N such that if $n \geqslant N$ it results
$\frac{1}{2\pi}\int_{\delta<|\theta|\leq\pi}|\phi_n(\theta)|d\theta \leq \epsilon$.

In order to illustrate the above definition, let us introduce the following simple and useful examples.

***Example** 2.2:* [21] The good examples of the 2pi-periodic approximate identity are the Jackson kernels

$$k_n(t) := k_{n,r}(t) := d_n\big(\frac{sin\frac{nt}{2}}{sin\frac{t}{2}}\big)^{2r}, r = 1, 2, ...$$

***Example** 2.3:* [22] Sinusoid function has also the 2pi-periodic approximate identity property.

We are going to present the definitions of the ϵ-net and the finite ϵ-net in the following. These definitions will be required for the proof of Theorem 3.2 in Section III and Theorem 4.3 in Section IV.

***Definition** 2.4:* [23] Let $\epsilon > 0$. A set $V_\epsilon \subset L^p_{2\pi}(\mathbb{R})$ is called ϵ-net of a set V, if $\tilde{f} \in V_\epsilon$ can be found for $\forall f \in V$ such that $\|f - \tilde{f}\|_{L^p_{2\pi}(\mathbb{R})} < \epsilon$.

***Definition** 2.5:* [23] The ϵ-net is said to be finite if it is a finite set of elements.

Here, we give the definition of the generalized Minkowski inequality which will be used for the proof of Theorem 4.2 in Section IV.

***Definition** 2.6:* [24] Let $h : [a, b] \times [c, d] \to \mathbb{R}$ be a integrable function, and $1 \leq p < \infty$, then the following inequality holds:

$$\Big(\int_a^b\Big|\int_c^d h(x,y)dy\Big|^p dx\Big)^{1/p} \leq \int_c^d\Big(\int_a^b|h(x,y)|^p dx\Big)^{1/p}dy$$

In the next section, we will derive some theoretical results concerning the universal approximation capability of a three-layer feedforward 2pi-periodic approximate identity neural networks in the space of continuous 2pi-periodic functions.

III. Theoretical results for continuous 2pi-periodic functions

In this section, we investigate the universal approximation capability of a three-layer feedforward 2pi-periodic approximate identity neural networks in the space of continuous 2pi-periodic functions. Let us begin with Theorem 3.1 which provides a connection between 2pi-periodic approximate identity and uniform convergence in the space of continuous 2pi-periodic functions. In fact, Theorem 3.1 shows any continuous 2pi-periodic function f converges to itself if it convolves with the 2pi-periodic approximate identity. Theorem 3.1 is required to present Theorem 3.2.

***Theorem** 3.1:* Let $C_{2\pi}(\mathbb{R})$ be a real linear space of all continuous 2pi-periodic functions on $\mathbb{R}$ to $\mathbb{R}$ with $\|f\|_{C_{2\pi}(\mathbb{R})} = \sup\ \{|f(x)|; |x| \leq \pi\}$. Let $\{\phi_n(\theta)\}_{n\in\mathbb{N}}$, $\phi_n(\theta) : [-\pi, \pi] \to \mathbb{R}$ be a 2pi-periodic approximate identity. Then for every $f\in C_{2\pi}(\mathbb{R})$, $\phi_n * f$ uniformly converges to f on $C_{2\pi}(\mathbb{R})$.

Proof: Let $\theta \in [-\pi, \pi]$ and $\epsilon > 0$. There exists a $\delta > 0$ such that $|f(\theta) - f(\zeta)| < \frac{\epsilon}{2\|\phi\|_{C_{2\pi}(\mathbb{R})}}$ for all $\zeta, |\theta - \zeta| < \delta$. Let us define $\{\phi_n * f\}_{n\in\mathbb{N}}$ by $\phi_n(\theta) = n\phi(n\theta)$, we consider

$$\begin{aligned} & \phi_n * f(\theta) - f(\theta) \\ = & \frac{1}{2\pi}\int_{-\pi}^{\pi} n\phi(n\xi)\{f(\theta-\xi) - f(\theta)\}d\xi \\ = & \frac{1}{2\pi}\Big(\int_{|\xi|<\delta} + \int_{\delta<|\xi|\leq\pi}\Big)n\phi(n\xi)\{f(\theta-\xi) - f(\theta)\}d\xi \\ = & I_1 + I_2. \end{aligned}$$

Subsequently we calculate I_1, I_2 as follows;

$$\begin{aligned} |I_1| &\leq \frac{1}{2\pi}\int_{|\xi|<\delta} n\phi(n\xi)\{f(\theta-\xi) - f(\theta)\}d\xi \\ &< \frac{\epsilon}{2\|\phi\|_{C_{2\pi}(\mathbb{R})}}\frac{1}{2\pi}\int_{|\xi|<\delta} n\phi(n\xi)d\xi \\ &= \frac{\epsilon}{2\|\phi\|_{C_{2\pi}(\mathbb{R})}}\frac{1}{2\pi}\int_{|\eta|<n\delta}\phi(\eta)d\eta \\ &\leq \frac{\epsilon}{2\|\phi\|_{C_{2\pi}(\mathbb{R})}}\frac{1}{2\pi}\int_{-\pi}^{\pi}\phi(\eta)d\eta = \frac{\epsilon}{2}. \end{aligned}$$

For I_2, we have

$$\begin{aligned}|I_2| &\leq 2\|f\|_{C_{2\pi}(\mathbb{R})}\frac{1}{2\pi}\int_{\delta<|\xi|\leq\pi} n|\phi(n\xi)|d\xi \\ &= 2\|f\|_{C_{2\pi}(\mathbb{R})}\frac{1}{2\pi}\int_{n\delta<|\eta|\leq n\pi} |\phi(\eta)|d\eta.\end{aligned}$$

Since

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{n\delta<|\eta|\leq n\pi}|\phi(\eta)|d\eta = 0,$$

there exists an $n_0 \in \mathbb{N}$ such that for all $n \geqslant n_0$,

$$\frac{1}{2\pi}\int_{n\delta<|\eta|\leq n\pi}|\phi(\eta)|d\eta < \frac{\epsilon}{4\|f\|_{C_{2\pi}(\mathbb{R})}}.$$

Combining I_1 and I_2 for $n \geqslant n_0$, we have

$$\|\phi_n * f(\theta) - f(\theta)\|_{C_{2\pi}(\mathbb{R})} < \epsilon.$$

■

The next theorem can be regarded as the main theorem of this section.

***Theorem** 3.2:* Let $C_{2\pi}(\mathbb{R})$ be a real linear space of all continuous 2pi-periodic functions on $\mathbb{R}$ to $\mathbb{R}$ with $\|f\|_{C_{2\pi}(\mathbb{R})} = \sup\{|f(x)|;|x| \leq \pi\}$, and $V \subset C_{2\pi}(\mathbb{R})$ a compact set. Let $\{\phi_n(\theta)\}_{n\in\mathbb{N}}$, $\phi_n : [-\pi,\pi] \to \mathbb{R}$ be a 2pi-periodic approximate identity. Let the family of functions $\{\sum_{j=1}^{M}\lambda_j\phi_j(\theta)|\lambda_j \in \mathbb{R}, \theta \in [-\pi,\pi], M \in \mathbb{N}\}$, be dense in $C_{2\pi}(\mathbb{R})$, and given $\epsilon > 0$. Then, there exists an $N \in \mathbb{N}$ which depends on V and ϵ but not on f, such that for any $f \in V$, there exist weights $c_k = c_k(f, V, \epsilon)$ satisfying

$$\left\|f(\theta) - \sum_{k=1}^{N} c_k\phi_k(\theta)\right\|_{C_{2\pi}(\mathbb{R})} < \epsilon.$$

Moreover, every c_k is a continuous function of $f \in V$.

Proof: Since V is compact, for any $\epsilon > 0$, there is a finite $\frac{\epsilon}{2}$-net $\{f^1, ..., f^M\}$ for V. This implies that for any $f \in V$, there is an f^j such that $\| f - f^j \|_{C_{2\pi}(\mathbb{R})} < \frac{\epsilon}{2}$. For any f^j, by assumption of the theorem, there are $\lambda_i^j \in \mathbb{R}, N_j \in \mathbb{N}$, and $\phi_i^j(\theta)$ such that

$$\left\|f^j(x) - \sum_{i=1}^{N_j}\lambda_i^j\phi_i^j(\theta)\right\|_{C_{2\pi}(\mathbb{R})} < \frac{\epsilon}{2}. \tag{1}$$

For any $f \in V$, we define

$$F_-(f) = \{j| \parallel f - f^j \parallel_{C_{2\pi}(\mathbb{R})} < \frac{\epsilon}{2}\},$$

$$F_0(f) = \{j| \parallel f - f^j \parallel_{C_{2\pi}(\mathbb{R})} = \frac{\epsilon}{2}\},$$

$$F_+(f) = \{j| \parallel f - f^j \parallel_{C_{2\pi}(\mathbb{R})} > \frac{\epsilon}{2}\}.$$

Therefore, $F_-(f)$ is not empty according to the definition of $\frac{\epsilon}{2}$-net. If $\widetilde{f} \in V$ approaches f such that $\| \widetilde{f} - f \|_{C_{2\pi}(\mathbb{R})}$ is small enough, then we have $F_-(f) \subset F_-(\widetilde{f})$ and $F_+(f) \subset F_+(\widetilde{f})$. Thus, $F_-(\widetilde{f}) \bigcap F_+(f) \subset F_-(\widetilde{f}) \bigcap F_+(\widetilde{f}) = \emptyset$, which implies $F_-(\widetilde{f}) \subset F_-(f) \cup F_0(f)$. We finish with the following.

$$F_-(f) \subset F_-(\widetilde{f}) \subset F_-(f) \cup F_0(f). \tag{2}$$

Define

$$d(f) = \left[\sum_{j\in F_-(f)}\left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{C_{2\pi}(\mathbb{R})}\right)\right]^{-1}$$

and

$$f_h = \sum_{j\in F_-(f)}\sum_{i=1}^{N_j} d(f)\left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{C_{2\pi}(\mathbb{R})}\right)\lambda_i^j\phi_i^j(\theta) \tag{3}$$

then, $f_h \in \left\{\sum_{j=1}^{M}\lambda_j\phi_j(\theta)\right\}$ approximates f with accuracy ϵ :

$$\begin{aligned}&\parallel f - f_h \parallel_{C_{2\pi}(\mathbb{R})} \\ = &\left\| \sum_{j\in F_-(f)} d(f)\left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{C_{2\pi}(\mathbb{R})}\right)\left(f - \sum_{i=1}^{N_j}\lambda_i^j\phi_i^j(\theta)\right)\right\|_{C_{2\pi}(\mathbb{R})} \\ = &\left\| \sum_{j\in F_-(f)} d(f)\left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{C_{2\pi}(\mathbb{R})}\right)\left(f - f^j + f^j - \sum_{i=1}^{N_j}\lambda_i^j\phi_i^j(\theta)\right)\right\|_{C_{2\pi}(\mathbb{R})} \\ \leq &\sum_{j\in F_-(f)} d(f)\left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{C_{2\pi}(\mathbb{R})}\right)\left(\|f - f^j \parallel_{C_{2\pi}(\mathbb{R})} + \left\|f_j - \sum_{i=1}^{N_j}\lambda_i^j\phi_i^j(\theta)\right\|_{C_{2\pi}(\mathbb{R})}\right) \\ \leq &\sum_{j\in F_-(f)} d(f)\left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{C_{2\pi}(\mathbb{R})}\right)\left(\frac{\epsilon}{2} + \frac{\epsilon}{2}\right) \\ = &\ \epsilon.\end{aligned} \tag{4}$$

In the next step, we prove the continuity of c_k. For the proof, we use (2) to obtain

$$\begin{aligned}&\sum_{j\in F_-(f)}\left(\frac{\epsilon}{2} - \parallel \widetilde{f} - f^j \parallel_{C_{2\pi}(\mathbb{R})}\right) \\ \leq &\sum_{j\in F_-(\widetilde{f})}\left(\frac{\epsilon}{2} - \parallel \widetilde{f} - f^j \parallel_{C_{2\pi}(\mathbb{R})}\right) \\ \leq &\sum_{j\in F_-(\widetilde{f})}\left(\frac{\epsilon}{2} - \parallel \widetilde{f} - f^j \parallel_{C_{2\pi}(\mathbb{R})}\right) + \\ &\sum_{j\in F_0(f)}\left(\frac{\epsilon}{2} - \parallel \widetilde{f} - f^j \parallel_{C_{2\pi}(\mathbb{R})}\right).\end{aligned} \tag{5}$$

Let $\widetilde{f} \to f$ in (5), then we have

$$\sum_{j \in F_-(\widetilde{f})} \left(\frac{\epsilon}{2} - \| \widetilde{f} - f^j \|_{C_{2\pi}(\mathbb{R})} \right) \to \sum_{j \in F_-(f)} \left(\frac{\epsilon}{2} - \| f - f^j \|_{C_{2\pi}(\mathbb{R})} \right). \tag{6}$$

This obviously demonstrates $d(\widetilde{f}) \to d(f)$. Thus, $\widetilde{f} \to f$ results

$$d(\widetilde{f}) \left(\frac{\epsilon}{2} - \| f - f^j \|_{C_{2\pi}(\mathbb{R})} \right) \lambda_i^j \to d(f) \left(\frac{\epsilon}{2} - \| f - f^j \|_{C_{2\pi}(\mathbb{R})} \right) \lambda_i^j. \tag{7}$$

Let $N = \sum_{j \in F_-(f)} N_j$ and define c_k in terms of

$$\begin{aligned} f_h &= \sum_{j \in F_-(f)} \sum_{i=1}^{N_j} d(f) \left(\frac{\epsilon}{2} - \| f - f^j \|_{C_{2\pi}(\mathbb{R})} \right) \lambda_i^j \phi_i^j(\theta) \\ &\equiv \sum_{k=1}^{N} c_k \phi_k(\theta) \end{aligned}$$

From (7), c_k is a continuous functional of f. Thus, the approximation result follows. ■

In the next section, we will obtain similar results in the space of pth-order Lebesgue integrable 2pi-periodic functions.

IV. Theoretical results for pth-order Lebesgue integrable 2pi-periodic functions

In this section, we investigate the universal approximate capability of a three-layer feedforward 2pi-periodic approximate identity neural networks in the space of pth-order Lebesgue-integrable 2pi-periodic functions. To prove Theorem 4.2, we need the following simple result.

***Theorem** 4.1:* [25] Assume $1 \leq p < \infty$. If $f \in L^p_{2\pi}(\mathbb{R})$, then $\lim_{y \to 0} \|f(x-y) - f(x)\|_{L^p_{2\pi}(\mathbb{R})} = 0$.

The following theorem provides a connection between 2pi-periodic approximate identity and uniform convergence in the space of pth-order Lebesgue-integrable 2pi-periodic functions. In fact, the following theorem shows that any pth-order Lebesgue-integrable 2pi-periodic function f converges to itself if it convolves with the 2pi-periodic approximate identity.

***Theorem** 4.2:* Let $L^p_{2\pi}(\mathbb{R})$ be a real linear space of all pth-order Lebesgue-integrable 2pi-periodic functions on $\mathbb{R}$ to $\mathbb{R}$ with $\|f\|_{L^p_{2\pi}(\mathbb{R})} = \{\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^p dx\}^{\frac{1}{p}}$. Let $\{\phi_n(\theta)\}_{n \in \mathbb{N}}$, $\phi_n(\theta) : [-\pi, \pi] \to \mathbb{R}$ be a 2pi-periodic approximate identity. Then for every $f \in L^p_{2\pi}(\mathbb{R})$, $\phi_n * f$ uniformly converges to f on $L^p_{2\pi}(\mathbb{R})$.

Proof: Generalized Minkowski inequality implies that

$$\begin{aligned} &\|\phi_n * f - f\|_{L^p_{2\pi}(\mathbb{R})} \\ \leq\ & \frac{1}{2\pi} \int_{-\pi}^{\pi} \|f(\theta - \xi) - f(\theta)\|_{L^p_{2\pi}(\mathbb{R})} |\phi_n(\xi)| d\xi. \end{aligned} \tag{8}$$

Using Theorem 4.1, for any $\epsilon > 0$ there exists a $\delta > 0$ such that if $|\xi| < \delta$

$$\|f(\theta - \xi) - f(\theta)\|_{L^p_{2\pi}(\mathbb{R})} \leq \frac{\epsilon}{2M} \tag{9}$$

Also the triangular inequality implies that

$$\|f(\theta - \xi) - f(\theta)\|_{L^p_{2\pi}(\mathbb{R})} \leq 2\|f\|_{L^p_{2\pi}(\mathbb{R})}. \tag{10}$$

By substituting the last two inequalities (9) and (10) in inequality (8), we obtain

$$\begin{aligned} &\|\phi_n * f - f\|_{L^p_{2\pi}(\mathbb{R})} \\ \leq\ & \frac{1}{2\pi} \int_{0<|\xi|\leq\delta} \frac{\epsilon}{2M} |\phi_n(\xi)| d\xi + \\ & \frac{1}{2\pi} \int_{\delta<|\xi|\leq\pi} 2\|f\|_{L^p_{2\pi}(\mathbb{R})} |\phi_n(\xi)| d\xi \\ \leq\ & \frac{\epsilon}{2M} \frac{1}{2\pi} \int_{0<|\xi|\leq\delta} |\phi_n(\xi)| d\xi + \\ & 2\|f\|_{L^p_{2\pi}(\mathbb{R})} \frac{1}{2\pi} \int_{\delta<|\xi|\leq\pi} |\phi_n(\xi)| d\xi \end{aligned} \tag{11}$$

By Definition 1, there exists an N such that for $n \geq N$

$$\frac{1}{2\pi} \int_{\delta<|\xi|\leq\pi} |\phi_n(\xi)| d\xi \leq \frac{\epsilon}{4\|f\|_{L^p_{2\pi}(\mathbb{R})}} \tag{12}$$

Using inequality (12) in (11), it follows that for $n \geq N$

$$\begin{aligned} &\|\phi_n * f - f\|_{L^p_{2\pi}(\mathbb{R})} \\ \leq\ & \frac{\epsilon}{2M}.M + 2\|f\|_{L^p_{2\pi}(\mathbb{R})} \frac{\epsilon}{4\|f\|_{L^p_{2\pi}(\mathbb{R})}} = \epsilon. \end{aligned}$$

■

Using the above theorem, we give the main result of this section in the following:

***Theorem** 4.3:* Let $L^p_{2\pi}(\mathbb{R})$ be a real linear space of all pth-order Lebesgue integrable 2pi-periodic functions on $\mathbb{R}$ to $\mathbb{R}$ with $\|f\|_{L^p_{2\pi}(\mathbb{R})} = \{\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^p dx\}^{\frac{1}{p}}$, and $V \subset L^p_{2\pi}(\mathbb{R})$ a compact set. Let $\{\phi_n(\theta)\}_{n \in \mathbb{N}}$, $\phi_n : [-\pi, \pi] \to \mathbb{R}$ be a 2pi -periodic approximate identity. Let the family of functions $\{\sum_{j=1}^{M} \lambda_j \phi_j(\theta) | \lambda_j \in \mathbb{R}, \theta \in [-\pi, \pi], M \in \mathbb{N}\}$, be dense in $L^p_{2\pi}(\mathbb{R})$, and given $\epsilon > 0$. Then, there exists an $N \in \mathbb{N}$ which depends on V and ϵ but not on f, such that for any $f \in V$, there exist weights $c_k = c_k(f, V, \epsilon)$ satisfying

$$\left\| f(\theta) - \sum_{k=1}^{N} c_k \phi_k(\theta) \right\|_{L^p_{2\pi}(\mathbb{R})} < \epsilon.$$

Moreover, every c_k is a continuous function of $f \in V$.

Proof: Since V is compact, for any $\epsilon > 0$, there is a finite $\frac{\epsilon}{2}$-net $\{f^1, ..., f^M\}$ for V. This implies that for any $f \in V$, there is an f^j such that $\| f - f^j \|_{L^p_{2\pi}(\mathbb{R})} < \frac{\epsilon}{2}$. For any f^j, by assumption of the theorem, there are $\lambda_i^j \in \mathbb{R}, N_j \in \mathbb{N}$, and $\phi_i^j(\theta)$ such that

$$\left\| f^j(x) - \sum_{i=1}^{N_j} \lambda_i^j \phi_i^j(\theta) \right\|_{L^p_{2\pi}(\mathbb{R})} < \frac{\epsilon}{2}. \quad (13)$$

For any $f \in V$, we define

$$F_-(f) = \{j | \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} < \frac{\epsilon}{2}\},$$

$$F_0(f) = \{j | \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} = \frac{\epsilon}{2}\},$$

$$F_+(f) = \{j | \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} > \frac{\epsilon}{2}\}.$$

Therefore, $F_-(f)$ is not empty according to the definition of $\frac{\epsilon}{2}$-net. If $\widetilde{f} \in V$ approaches f such that $\| \widetilde{f} - f \|_{L^p_{2\pi}(\mathbb{R})}$ is small enough, then we have $F_-(f) \subset F_-(\widetilde{f})$ and $F_+(f) \subset F_+(\widetilde{f})$. Thus, $F_-(\widetilde{f}) \bigcap F_+(f) \subset F_-(\widetilde{f}) \bigcap F_+(\widetilde{f}) = \emptyset$, which implies $F_-(\widetilde{f}) \subset F_-(f) \cup F_0(f)$. We finish with the following.

$$F_-(f) \subset F_-(\widetilde{f}) \subset F_-(f) \cup F_0(f). \quad (14)$$

Define

$$d(f) = \left[\sum_{j \in F_-(f)} \left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) \right]^{-1}$$

and

$$f_h = \sum_{j \in F_-(f)} \sum_{i=1}^{N_j} d(f) \left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) \lambda_i^j \phi_i^j(\theta) \quad (15)$$

then, $f_h \in \left\{ \sum_{j=1}^{M} \lambda_j \phi_j(\theta) \right\}$ approximates f with accuracy ϵ :

$$\begin{aligned}
& \parallel f - f_h \parallel_{L^p_{2\pi}(\mathbb{R})} \\
= & \left\| \sum_{j \in F_-(f)} d(f) \left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) \right. \\
& \left. \left(f - \sum_{i=1}^{N_j} \lambda_i^j \phi_i^j(\theta) \right) \right\|_{L^p_{2\pi}(\mathbb{R})} \\
= & \left\| \sum_{j \in F_-(f)} d(f) \left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) \right. \\
& \left. \left(f - f^j + f^j - \sum_{i=1}^{N_j} \lambda_i^j \phi_i^j(\theta) \right) \right\|_{L^p_{2\pi}(\mathbb{R})} \\
\leq & \sum_{j \in F_-(f)} d(f) \left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) \\
& \left(\| f - f^j \|_{L^p_{2\pi}(\mathbb{R})} + \left\| f_j - \sum_{i=1}^{N_j} \lambda_i^j \phi_i^j(\theta) \right\|_{L^p_{2\pi}(\mathbb{R})} \right) \\
\leq & \sum_{j \in F_-(f)} d(f) \left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) \left(\frac{\epsilon}{2} + \frac{\epsilon}{2} \right) \\
= & \epsilon. \qquad (16)
\end{aligned}$$

In the next step, we prove the continuity of c_k. For the proof, we use (14) to obtain

$$\begin{aligned}
& \sum_{j \in F_-(f)} \left(\frac{\epsilon}{2} - \parallel \widetilde{f} - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) \\
\leq & \sum_{j \in F_-(\widetilde{f})} \left(\frac{\epsilon}{2} - \parallel \widetilde{f} - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) \\
\leq & \sum_{j \in F_-(\widetilde{f})} \left(\frac{\epsilon}{2} - \parallel \widetilde{f} - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) + \\
& \sum_{j \in F_0(f)} \left(\frac{\epsilon}{2} - \parallel \widetilde{f} - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right). \qquad (17)
\end{aligned}$$

Let $\widetilde{f} \to f$ in (17), then we have

$$\begin{aligned}
& \sum_{j \in F_-(\widetilde{f})} \left(\frac{\epsilon}{2} - \parallel \widetilde{f} - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) \to \\
& \sum_{j \in F_-(f)} \left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right). \qquad (18)
\end{aligned}$$

This obviously demonstrates $d(\widetilde{f}) \to d(f)$. Thus, $\widetilde{f} \to f$ results

$$\begin{aligned}
& d(\widetilde{f}) \left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) \lambda_i^j \to \\
& d(f) \left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) \lambda_i^j. \qquad (19)
\end{aligned}$$

Let $N = \sum_{j \in F_-(f)} N_j$ and define c_k in terms of

$$\begin{aligned}
f_h &= \sum_{j \in F_-(f)} \sum_{i=1}^{N_j} d(f) \left(\frac{\epsilon}{2} - \parallel f - f^j \parallel_{L^p_{2\pi}(\mathbb{R})} \right) \lambda_i^j \phi_i^j(\theta) \\
&\equiv \sum_{k=1}^{N} c_k \phi_k(\theta)
\end{aligned}$$

From (19), c_k is a continuous functional of f. Thus, the approximation result follows. ■

V. Conclusions

In this study, we have investigated the universal approximation capabilities of a three-layer feedforward 2pi-periodic approximate identity neural networks. We have reviewed the basic definition of the 2pi-periodic approximate identity. Using 2pi-periodic approximate identity, we have presented Theorem 3.1 that shows any continuous 2pi-periodic function f converges to itself if it convolves with the 2pi-periodic approximate identity. Using Theorem 3.1, we have also presented Theorem 3.2 that shows three-layer feedforward 2pi-periodic approximate identity neural networks are universal approximators in the space of continuous 2pi-periodic functions. Making use of 2pi-periodic approximate identity again, we have proved Theorem 4.2 that shows any pth-order Lebesgue-integrable 2pi-periodic function f converges to itself if it convolves with the 2pi-periodic approximate identity. Using Theorem 4.2, we have also proved Theorem 4.3 that provides an explanation of how can the universal approximation capability of these networks occur in the space of pth-order Lebesgue integrable 2pi-periodic functions. In fact, The above theoretical results are useful in understanding the approximation theory of a three-layer feed forward 2pi-periodic approximate identity neural networks by using the theory of ϵ-net.

ACKNOWLEDGMENTS

We thank the anonymous referees for their valuable suggestions. This work was supported by Universiti Sains Malaysia (Grant No. 1001/PMATHS/811161).

REFERENCES

[1] V. E., Ismailov, "Approximation by neural networks with weights varying on a finite set of directions," *Journal of Mathematical Analysis and Applications*, Vol. 389(1): 72-83, 2010.

[2] G., Cybenko, " Approximation by superposition of sigmoidal functions," *Math Control Signals Systems*, Vol. 2(4): 303-314, 1989.

[3] J., Park, I. W., Sandberg, " Universal approximation using radial-basis-function networks," *Neural Computation*, Vol. 3(2): 246-257, 1991.

[4] F., Scarcelli, A. C., Tsoi "Universal approximation using feedforward neural networks: A survey of some existing methods and some new results," *Neural Networks*, Vol. 11(1): 15-37, 1998.

[5] M., Sanguineti, " Universal approximation by ridge Computational models and neural networks: a survey," *The Open Applied Mathematics Journal*, Vol. 2: 31-58, 2008.

[6] X., Ting-f., Z., Xin-l., " Neural networks for optimal approximation of continuous function in $\mathbb{R}^d$," *Appl. Math. J. Chinese Univ.*, Vol. 27(3): 335-344, 2012.

[7] D. S. Yu, " Approximation by neural networks with sigmoidal functions," *Acta Mathematica Sincica*, Vol. 29(10): 2013-2026, 2013.

[8] B. Lenze, " Note on a density question for neural networks," *Numerical Functional Analysis and Optimization*, Vol. 15(7-8): 909-913, 1994.

[9] N., Hahm, B. I., Hong, " The capability of periodic neural networks approximation," *Korean J. Math*, Vol. 18(2): 167-174, 2010.

[10] G. S. da S., Gomes, T. B., Ludermir, L. M. M. R., Lima, "Comparison of new activation functions in neural network for forecasting financial time series," *Neural Computing and Applications*, Vol. 20(3): 417-439, 2011.

[11] J., Wang, B., Chen, C., Yang, " Approximation of algebraic and trigonometric polynomials by feedforward neural networks," *Neural Computing and Applications*, Vol. 21(1): 73-80, 2012.

[12] Z. X., Chen, F. L., Cao, J. W., Zhao, "The construction and approximation of some neural networks operators," *Appl. Math. J. Chinese Univ.*, Vol. 27(1): 69-77, 2012.

[13] C. Turchetti, M. Conti, P. Crippa, S. Orcioni, "On the approximation of stochastic processes by approximate identity neural networks," *IEEE Transaction on Neural Networks*, Vol. 9(6): 1069-1085, 1998.

[14] Z., Zainuddin, S., Panahian Fard, " Double approximate identity neural networks universal approximation in real Lebesgue spaces," *Lecture Notes in Computer Science*, Vol. 7663: 409-415, 2012.

[15] S., Panahian Fard, Z., Zainuddin, " On the universal approximation capability of flexible approximate identity neural networks," *Lecture Notes in Electrical Engineering*, Vol. 236: 201-207, 2013.

[16] S., Panahian Fard, Z., Zainuddin, " Analyses for $L^p[a,b]$-norm approximation capability of flexible approximate identity neural networks," *Neural Computing and Applications*, DOI 10.1007/s00521-013-1493-9, 2013.

[17] S., Panahian Fard, Z., Zainuddin, "The universal approximation capabilities of Mellin approximate identity neural networks," *Lecture Notes in Computer Science*, Vol. 7951: 205-213, 2013.

[18] S., Panahian Fard, Z., Zainuddin, "The universal approximation capability of double flexible approximate identity neural networks," *Lecture Notes in Electrical Engineering*, Vol. 277: DOI 10.1007/978-3-319-01766-2-15, 2014.

[19] W., Wu, D., Nan, Z., Li, J., Long, " Approximation to compact set of functions by feedforward neural networks," *In 20th International Joint Conference on Neural Networks*, Orlando, Fl, USA, pp. 1222-1225, 2007.

[20] B. R. Draganov, "On the approximation by convolution operators in homogeneous banach spaces of periodic functions," *Mathematica Balkanica*, Vol. 25(1-2): 39-59, 2011.

[21] A. R., Devore, D., Leviatan, X. M., Yu, "L_p approximation by reciprocals of trigonometric and algebraic polynomials," *Canad. Math. Bull*, Vol. 33(4): 460-469, 1990.

[22] A., Ismail, D.-S., Jeng, L. L., Zhang, J. S., Zhang, " Predictions of bridge scour: Application of a feed-forward neural network with an adaptive activation function," *Engineering Applications of Artificial Intelligence*, Vol. 26(5-6): 1540-1549, 2013.

[23] V., Lebedev, " An introduction to functional analysis and computational mathematics," *Brikhäuser*, Boston, 1997, pp.16.

[24] A., Zygmund, " Trigonometric series," *Cambridge University Press*, New York, 1968, pp. 19.

[25] F., Jones, " Lebesgue integration on Euclidean space, Jones and Bartlett, Boston, 1993, pp. 245.