J. Math. Computer Sci., 38 (2025), 111–124

Online: ISSN 2008-949X

Journal of Mathematics and Computer Science

Journal Homepage: www.isr-publications.com/jmcs

A double inertial embedded Ishikawa algorithm for two nonexpansive mappings applies to breast cancer detection

Pennipat Nabheerong[a], Warissara Kiththiworaphongkich[b], Watcharaporn Cholamjiak[c,*]

[a]Department of Radiology, School of Medicine, University of Phayao, Phayao 56000, Thailand.

[b]Department of Radiology, Phayao Hospital, Phayao 56000, Thailand.

[c]Department of Mathematics, School of Science, University of Phayao, Phayao 56000, Thailand.

Abstract

Breast cancer remains a major health challenge worldwide, with early detection being crucial for enhancing treatment outcomes and prolonging life. This study introduces innovative machine learning techniques aimed at advancing breast cancer screening models. Specifically, we propose a double inertial embedded Ishikawa algorithm that enhances the traditional Ishikawa algorithm with a double inertial technique for two distinct nonexpansive mappings. We establish a weak convergence theorem in Hilbert spaces, utilizing relaxed real number extrapolation parameters. An example in an infinite-dimensional space corroborates our theoretical findings. The practical application of this algorithm is demonstrated using a real dataset from Phayao Hospital in Northern Thailand, supplemented by data from the UCI website, within an Extreme Learning Machine framework for efficient prediction. Achieving an accuracy of 88.57%, our algorithm outperforms the well-known FISTA algorithms, highlighting its potential for improving breast cancer diagnostic protocols and outcomes.

Keywords: Nonexpansive, Ishikawa algorithm, inertial technique, mammographic, breast cancer.

2020 MSC: 46E20, 52A07, 68Q04.

1. Introduction

In this paper, we consistently assume that H is a real Hilbert space equipped with the norm $\|\cdot\|$ and the inner product $\langle\cdot,\cdot\rangle$, $T : H \to H$ is a nonexpansive mapping, if $\|Ta - Tb\| \leqslant \|a - b\|$ for all $a, b \in H$, $F(T)$ is denoted the set of fixed point set of T, i.e., $F(T) = \{a^* \in H : a^* = Ta^*\}$. Solving fixed point problems of nonexpansive mappings is widely used in science and engineering since many issues in these fields can be reformulated into nonexpansive transmission equations, see in [4, 5]. The existence results for nonexpansive mappings in different completed metric spaces was first obtained by Browder [9] and Kirk [26] in 1965, independently. The different way to study nonexpansive mapping is fixed point approximation by modifying many iterative algorithms, see in [1, 10, 44]. One of the famous algorithms is well known as Mann's algorithm, which Mann [36] introduced. The algorithm was generated by $a^0 \in H$ and

$$a^{k+1} = (1 - \alpha^k)a^k + \alpha^k Ta^k, \ \forall k \in \mathbb{N},$$

where α^k is a nonnegative sequence in $[0, 1]$. Another famous algorithm is also well known as Ishikawa's

*Corresponding author

Email addresses: pennipat.na@up.ac.th (Pennipat Nabheerong), Bow_wk@hotmail.com (Warissara Kiththiworaphongkich), watcharaporn.ch@up.ac.th (Watcharaporn Cholamjiak)

doi: 10.22436/jmcs.038.01.08

Received: 2024-07-15 *Revised:* 2024-08-27 *Accepted:* 2024-10-11

algorithm which was introduced by Ishikawa [20]. The algorithm was generated by $a^0 \in H$ and

$$a^{k+1} = (1-\beta^k)a^k + a^k T((1-\alpha^k)a^k + \alpha^k T a^k),\ \forall k \in \mathbb{N},$$

where α^k and β^k are nonnegative sequences in $[0,1]$. In 1990, Goebel and Kirk [16] proved the useful lemma called the demiclosed principle of a nonexpansive T, that is if there exists a sequence $\{a^k\}$ in H such that $a^k \rightharpoonup a^* \in H$ and $\|a^k - Ta^k\| \rightharpoonup 0$, then $a^* \in F(T)$. To accelerate the convergence of algorithms, the inertial extrapolation technique is the one efficient technique which is also called the heavy ball method proposed by Polyak [43] in 1964. The algorithm was generated by $a^0, a^1 \in H$, $\gamma > 0$, and

$$a^{k+1} = a^k + \theta^k(a^k - a^{k-1}) - \gamma^k \nabla f(a^k),\ \forall k \in \mathbb{N},$$

where $f : H \to H$ is differentiable and $\theta^k \in [0,1)$ is the extrapolation coefficient of the inertial step $\theta^k(a^k - a^{k-1})$. Later, the inertial technique was used to modify the algorithm to speed up the convergence of the algorithms by many mathematicians [2, 8, 34, 42]. Efficient and popular algorithm modified by the inertial technique [30] is fast iterative shrinkage-thresholding algorithm (for short FISTA) which was introduced by Beck and Teboulle [6] for minimization problems. The FISTA algorithm is in form as follows:

$$\begin{cases} b^1 = a^0 \in H, t^1 = 1, \\ a^k = \text{prox}_{\frac{1}{L}g}(b^k - \frac{1}{L}\nabla f(b^k)), \\ t^{k+1} = \frac{1+\sqrt{1+4(t^k)^2}}{2}, \theta_{\text{FISTA}} = \frac{t^{k-1}}{t^{k+1}}, \\ b^{k+1} = a^k + \theta_{\text{FISTA}}(a^k - a^{k-1}), \end{cases}$$

where $f : H \to H$ is convex with ∇f is Lipschitz continuous with constant L and $g : H \to H$ is closed and convex. It's well known that the mapping $\text{prox}_{\gamma g}(I - \gamma\nabla f(\cdot))$ is nonexpansive when $\gamma \in (0, \frac{2}{L}]$ [48].

Recently, Nabheerong et al. [39] introduced the inertial relaxed CQ algorithm with Mann's iteration to address split feasibility problems in breast cancer detection, utilizing the mammographic mass dataset from the UCI repository. The proposed algorithm effectively classified the binary severity of breast cancer based on a well-fitted model. A more complex challenge beyond binary classification is predicting the seven BI-RADS (Breast Imaging-Reporting and Data System) categories, a widely accepted risk assessment and quality assurance tool developed by the American College of Radiology for breast imaging interpretation.

Building on these findings, we propose a double inertial embedded Ishikawa algorithm for two nonexpansive mappings aimed at breast cancer detection, designed to approximate a common fixed point of two nonexpansive mappings in real Hilbert spaces. Furthermore, as tumor invasion models are frequently governed by chemotaxis mechanisms represented by partial differential equations, incorporating details in this direction would offer a more comprehensive perspective to the reader; relevant examples include [12, 23, 24, 28, 29]. We shall apply our algorithm for a real data set from Phayao Hospital, northern Thailand. Our paper is organized as follows. In Section 2, we present some preliminary and mathematical tools. In Section 3, we propose a double inertial embedding in the Ishikawa algorithm and prove the weak convergence results under certain assumptions. Section 4 contains applications using our proposed method and compares it with other related methods in the literature.

2. Preliminaries

In this section, let H be a real Hilbert space and C be a closed convex subset of H. We denote the weak and strong convergence of a sequence a^k to a point $a^* \in H$ as $a^k \rightharpoonup a^*$ and $a^k \to a^*$, respectively. We are listing some crucial results necessary to prove our main result. Recall that a mapping $T : H \to H$ is said to be L-Lipschitz continuous if there is a constant $L > 0$ such that

$$\|Ta - Tb\| \leqslant L\|a - b\|,\ \forall a, b \in H.$$

If $L = 1$, then T is said to be nonexpansive.

Lemma 2.1 ([16]). *Let $T : H \to H$ be a nonexpansive mapping such that $F(T) \neq \emptyset$. If there exists a sequence $\{a^k\}$*

in H *such that* $a^k \rightharpoonup a^* \in H$ *and* $\|a^k - Ta^k\| \to 0$, *then* $x \in F(T)$.

Lemma 2.2 ([41]). *Let* C *be a nonempty set of* H *and* $\{a^k\}$ *be a sequence in* H. *Assume that the following conditions hold.*

(i) *For every* $a^* \in C$, *the sequence* $\{\|a^k - a^*\|\}$ *converges.*
(ii) *Every weak sequential cluster point of* $\{a^k\}$ *belongs to* C.

Then $\{a^k\}$ *weakly converges to a point in* C.

Lemma 2.3 ([40]). *Suppose that* $\{\gamma^k\}$, $\{\xi^k\}$, *and* $\{\upsilon^k\}$ *are sequences in* $[0, +\infty)$ *such that* $\gamma^{k+1} \leqslant \gamma^k + \upsilon^k(\gamma^k - \gamma^{k-1}) + \xi^k$, $\forall k \geqslant 1$, $\sum_{k=1}^{\infty} \xi^k < +\infty$ *and there is* $\upsilon \in \mathbb{R}$ *with* $0 \leqslant \upsilon^k < \upsilon < 1$, $\forall k \geqslant 1$. *Then the following conditions are satisfied:*

(i) $\sum [\gamma^k - \gamma^{k-1}]_+ < +\infty$, *where* $[r]_+ = \max\{r, 0\}$;
(ii) *there exists* $\gamma^* \in [0, \infty)$ *such that* $\lim_{k\to+\infty} \gamma^k = \gamma^*$.

3. Main results

Throughout the paper, we suppose that C is a nonempty closed and convex subset of a real Hilbert space H. Let $T_1, T_2 : H \to H$ be nonexpansive mappings such that $F(T_1) \cap F(T_2) \neq \emptyset$.

Algorithm 3.1 (Double inertial embedded Ishikawa algorithm)**.**

Initialization: Select arbitrary points $b^{-1}, b^0, a^0 \in H, \{\alpha^k\}, \{\beta^k\} \subset (a, b) \subset (0, 1)$, and $\{\theta^k\}, \{\delta^k\} \subset (-\infty, \infty)$. Set $k = 0$

Iterative Steps: Compute a^{k+1} as follows.

Step 1. Calculate $b^{k+1} = (1 - \alpha^k)a^k + \alpha^k T_1 a^k$.
Step 2. Calculate $c^{k+1} = b^{k+1} + \theta^k(b^{k+1} - b^k) + \delta^k(b^k - b^{k-1})$.
Step 3. Calculate $a^{k+1} = (1 - \beta^k)c^{k+1} + \beta^k T_2 c^{k+1}$. Set $k := k + 1$ and go to Step 1.

Theorem 3.2. *The sequence* $\{a^k\}$ *generated by Algorithm 3.1 weakly converges to an element of* $F(T_1) \cap F(T_2)$. *Assume that the following conditions hold:*

$$\sum_{k=1}^{\infty} |\theta^k| \|b^{k+1} - b^k\| < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} |\delta^k| \|b^k - b^{k-1}\| < \infty.$$

Proof. Let $p \in F(T_1) \cap F(T_2)$. By the nonexpansiveness of T_1, T_2, we have

$$\begin{aligned}
\|a^{k+1} - p\| &= \|(1 - \beta^k)c^{k+1} + \beta^k T_2 c^{k+1} - p\| \\
&\leqslant (1 - \beta^k)\|c^{k+1} - p\| + \beta^k \|T_2 c^{k+1} - p\| \\
&\leqslant \|c^{k+1} - p\| \\
&\leqslant \|b^{k+1} - p\| + |\theta^k| \|b^{k+1} - b^k\| + |\delta^k| \|b^k - b^{k-1}\| \\
&\leqslant (1 - \alpha^k)\|a^k - p\| + \alpha^k \|T_1 a^k - p\| + |\theta^k| \|b^{k+1} - b^k\| + |\delta^k| \|b^k - b^{k-1}\| \\
&\leqslant \|a^k - p\| + |\theta^k| \|b^{k+1} - b^k\| + |\delta^k| \|b^k - b^{k-1}\|.
\end{aligned}$$

By our conditions, it follows from Lemma 2.3 that $\lim_{k\to\infty} \|a^k - p\|$ exists. This implies that $\{a^k\}$ is bounded. It's also $\{b^k\}$ bounded. On the other hand, we have

$$\begin{aligned}
\|a^{k+1} - p\|^2 &= (1 - \beta^k)\|c^{k+1} - p\|^2 + \beta^k \|T_2 c^{k+1} - p\|^2 - (1 - \beta^k)\beta^k \|T_2 c^{k+1} - c^{k+1}\|^2 \\
&\leqslant \|c^{k+1} - p\|^2 - (1 - \beta^k)\beta^k \|T_2 c^{k+1} - c^{k+1}\|^2
\end{aligned}$$

$$\begin{aligned}
&\leqslant \|b^{k+1}-p\|^2+2\langle\theta^k(b^{k+1}-b^k)+\delta^k(b^k-b^{k-1}),c^{k+1}-p\rangle\\
&\quad-(1-\beta^k)\beta^k\|T_2c^{k+1}-c^{k+1}\|^2 \qquad (3.1)\\
&=(1-\alpha^k)\|a^k-p\|^2+\alpha^k\|T_1a^k-p\|^2-(1-\alpha^k)\alpha^k\|T_1a^k-a^k\|^2\\
&\quad+2\langle\theta^k(b^{k+1}-b^k)+\delta^k(b^k-b^{k-1}),c^{k+1}-p\rangle-(1-\beta^k)\beta^k\|T_2c^{k+1}-c^{k+1}\|^2\\
&\leqslant \|a^k-p\|^2-(1-\alpha^k)\alpha^k\|T_1a^k-a^k\|^2-(1-\beta^k)\beta^k\|T_2c^{k+1}-c^{k+1}\|^2\\
&\quad+2\langle\theta^k(b^{k+1}-b^k)+\delta^k(b^k-b^{k-1}),c^{k+1}-p\rangle.
\end{aligned}$$

Again by our conditions and (3.1), we obtain

$$\lim_{k\to\infty}\|T_1a^k-a^k\|=\lim_{k\to\infty}\|T_2c^{k+1}-c^{k+1}\|=0. \qquad (3.2)$$

By the definition of $\{b^{k+1}\}$, we have

$$\lim_{k\to\infty}\|b^{k+1}-a^k\|=\lim_{k\to\infty}\alpha^k\|T_1a^k-a^k\|=0. \qquad (3.3)$$

By the definition of $\{c^{k+1}\}$, we also have

$$\lim_{k\to\infty}\|c^{k+1}-b^{k+1}\|\leqslant\lim_{k\to\infty}(|\theta^k|\|b^{k+1}-b^k\|+|\delta^k|\|b^k-b^{k-1}\|)=0. \qquad (3.4)$$

It follows from (3.3) and (3.4) that

$$\lim_{k\to\infty}\|c^{k+1}-a^k\|=0. \qquad (3.5)$$

Since $\{a^k\}$ is bounded, suppose that a^* is a weak sequential cluster point of a^k. It follows by the equation (3.5) that a^* is also a weak sequential cluster point of c^{k+1}. By using Lemma 2.1 with (3.2), we can get that $a^*\in F(T_1)\cap F(T_2)$. By applying opial's lemma (Lemma 2.2), we can obtain that $\{a^k\}$ converges weakly to an element in $F(T_1)\cap F(T_2)$. □

Remark 3.3.

(i) For applying to solve the variational inclusion problem (VIP) which is to find $a^*\in H$ such that

$$0\in(Fx+Gx), \qquad (3.6)$$

where $F:H\to H$ is τ-inverse strongly monotone mapping and $G:H\to 2^H$ is a maximal monotone mapping. It is well known that $\gamma\in(0,2\tau)$, then the mapping $J_\gamma^G(I-\gamma F)$ is nonexpansive where J_γ^G is the resolvent of the operator G and $F(J_\gamma^G(I-\gamma F))=(F+G)^{-1}(0)$ which is the solution set of the VIP [33], respectively.

(ii) For applying the metric projection P_C to constrain the solution on a closed convex set C, we can set $T_1\equiv P_C$ or $T_2\equiv P_C$, then $C=F(T_2)$ or $C=F(T_1)$.

From the above remark, we can modify two different double inertial forward-backward splitting algorithms converging weakly to an element in $(F+G)^{-1}(0)\cap C$ by the same proof of Theorem 3.2. These algorithms are modified as follows.

Setting $T_1\equiv J_\gamma^G(I-\gamma F)$ and $T_2\equiv P_C$, we have following.

Algorithm 3.4 (Double inertial projection forward-backward splitting algorithm 1)**.**

Initialization: Select arbitrary points $b^{-1},b^0,a^0\in H,\{\alpha^k\},\{\beta^k\}\subset(a,b)\subset(0,1)$, and $\{\theta^k\},\{\delta^k\}\subset(-\infty,\infty)$. Set $k=0$

Iterative steps: Compute a^{k+1} as follows.

Step 1. Calculate $b^{k+1}=(1-\alpha^k)a^k+\alpha^kJ_\gamma^G(I-\gamma F)a^k$.

Step 2. Calculate $c^{k+1}=b^{k+1}+\theta^k(b^{k+1}-b^k)+\delta^k(b^k-b^{k-1})$.

Step 3. Calculate $a^{k+1} = (1-\beta^k)c^{k+1} + \beta^k P_C c^{k+1}$. Set $k := k+1$ and go to Step 1.

Setting $T_1 \equiv P_C$ and $T_2 \equiv J_\gamma^G(I-\gamma F)$, we have following.

Algorithm 3.5 (Double inertial projection forward-backward splitting algorithm 2)**.**

Initialization: Select arbitrary points $b^{-1}, b^0, a^0 \in H, \{\alpha^k\}, \{\beta^k\} \subset (a,b) \subset (0,1)$, and $\{\theta^k\}, \{\delta^k\} \subset (-\infty, \infty)$. Set $k = 0$

Iterative Steps: Compute a^{k+1} as follows.

Step 1. Calculate $b^{k+1} = (1-\alpha^k)a^k + \alpha^k P_C a^k$.
Step 2. Calculate $c^{k+1} = b^{k+1} + \theta^k(b^{k+1} - b^k) + \delta^k(b^k - b^{k-1})$.
Step 3. Calculate $a^{k+1} = (1-\beta^k)c^{k+1} + \beta^k J_\gamma^G(I-\gamma F)c^{k+1}$. Set $k := k+1$ and go to Step 1.

We shall give an example in the infinitely dimension space $L_2[0,1] = \{s(t) : \int_0^1 s(t)dt < \infty\}$, where such that $\|\cdot\|$ is L_2-norm defined by $\|s\| = \sqrt{\int_0^1 |s(t)|^2 dt}$ for supporting our main theorem (Algorithm 3.1).

Example 3.6. Let $H = L_2[0,1]$, $T_1 s(t) = \frac{3}{4}s(t)$ and $T_2 s(t) = \frac{2}{3}s(t)$ where $s(t) \in L_2[0,1]$. In our experiments, we use the Cauchy errors $\|a^k - a^{k+1}\|^2 < 10^{-4}$ to stop the iteration. To meet the highest performance of the our algorithm, we shall consider the necessary parameters of Algorithm 3.1 in 3 cases when

$$\theta^k = \begin{cases} \frac{1}{\|a^k - a^{k-1}\|^3}, & \text{if } a^k \neq a^{k-1} \text{ and } k > M, \\ \theta, & \text{otherwise,} \end{cases} \tag{3.7}$$

and

$$\delta^k = \begin{cases} \frac{1}{\|a^{k-1} - a^{k-2}\|^3}, & \text{if } a^{k-1} \neq a^{k-2} \text{ and } k > M, \\ \delta, & \text{otherwise,} \end{cases} \tag{3.8}$$

where M is the iteration number that we want to stop, and we choose the initialization $a^0 = e^t - 2$, $b^{-1} = \sin t$, and $b^0 = a^0$.

Case1. Set $\theta = \theta_{\text{FISTA}}$, $\delta = -0.001$, and $\beta^k = 0.5$. The different parameters α^k are considered in Table 1.

Table 1: Numerical results of different parameters α^k.

α^k	$\frac{k}{1.1k+1}$	$\frac{k}{5k+1}$	$\frac{k}{10k+1}$	$\frac{k}{100k+1}$	$\frac{k}{10^4k+1}$
CPU Time	0.9338	1.1016	1.1606	1.2240	1.2613
Number of iterations	12	16	17	18	18

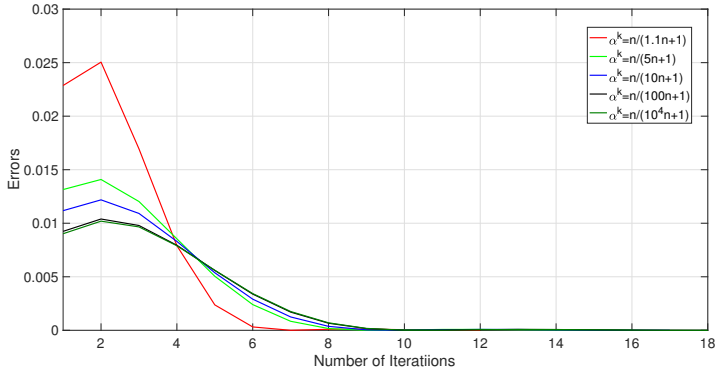

Figure 1: Cauchy errors plots of the different parameters α^k used in Algorithm 3.1 for Example 3.6.

Case2. Set $\theta = \theta_{FISTA}$, $\delta = -0.001$, and $\alpha^k = \frac{k}{1.1k+1}$. The different parameters β^k are considered in Table 2.

Table 2: Numerical results of different parameters β^k.

β^k	$\frac{k}{1.1k+1}$	$\frac{k}{5k+1}$	$\frac{k}{10k+1}$	$\frac{k}{100k+1}$	$\frac{k}{10^4k+1}$
CPU Time	0.9766	1.1589	1.1881	0.8457	1.3010
Number of iterations	13	16	17	11	19

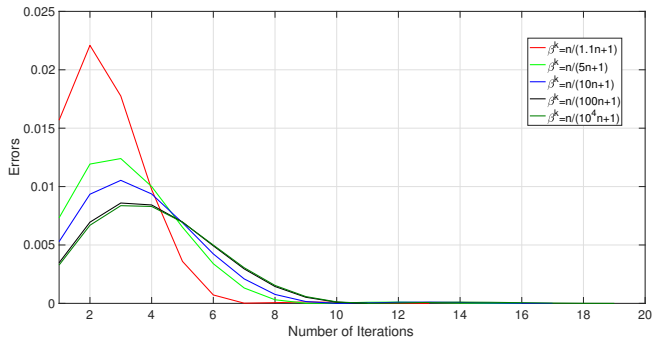

Figure 2: Cauchy errors plots of the different parameters β^k used in Algorithm 3.1 for Example 3.6.

Case3 Set $\alpha^k = \frac{k}{1.1k+1}$ and $\beta^k = \frac{k}{100k+1}$. The different parameters θ and δ are presented in Table 3.

Table 3: Numerical results of different parameters θ and δ.

	$\theta = 0, \delta = -0.001$	$\theta = 0, \delta = 0$	$\theta = \theta_{FISTA}, \delta = -0.001$
CPU Time	1.7037	1.7959	0.8169
Number of iterations	28	28	11

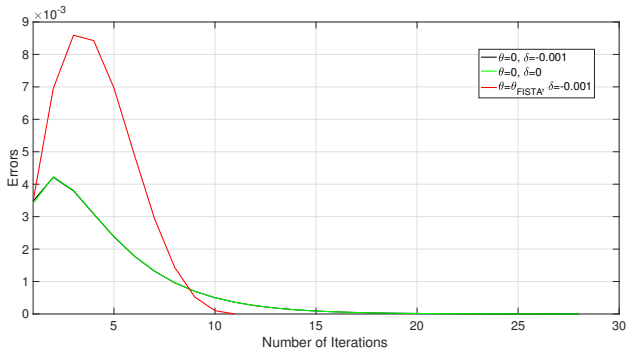

Figure 3: Cauchy errors plots of the different parameters θ and δ used in Algorithm 3.1 for Example 3.6.

Analysis of Tables 1-3 and Figures 1-3 indicates that the parameters $\alpha^k = \frac{k}{1.1k+1}$, $\beta^k = \frac{k}{100k+1}$, $\theta = \theta_{FISTA}$, and $\delta = -0.001$ yield the most favorable results for the initial conditions $a^0 = e^t - 2$, $b^{-1} = \sin t$, and $b^0 = a^0$.

4. Application

Breast cancer is the most common cancer in women, with approximately two million new cases of breast cancer diagnosed each year worldwide [18, 35]. There has been an increase in death rates over the past decade due to various factors such as the world entering an aging society [35]. Medical advances and an effective screening system makes it possible to detect breast cancer early. Currently, there are many studies conclude that screening for breast cancer with mammography in women aged 40 years and older can reduce the death rate [17, 25, 37], increase survival rate and helps planning treatment from the

beginning [13]. However, testing by this method has been found to have some disadvantages, such as over diagnosis. It can make the patient more anxious, more expose to radiation, including unnecessary biopsy. Moreover, similarities between early signs of breast cancer and normal structures in these images make detection and diagnosis of breast cancer a difficult task [3]. Diagnosis or screening using artificial intelligence (AI) or computer-aided diagnosis (CAD) is therefore another interesting modality. Researchers have studied the accuracy of AI in detecting breast cancer in mammography images, believing that using AI can help radiologists be more accurate in diagnosing breast cancer [11, 27]. As for CAD, it is considered to play a role in helping diagnose breast cancer as well. Computer interpretation provides a second opinion, helping radiologists review and make more confident interpretations of mammography results [21, 22, 25, 32, 45]. However, some studies have found that the role of AI in this is unclear [15]. It is seen that every sector should study or develop a form of AI in order to understand and be able to use it to help diagnose breast cancer as accurately as possible.

In 2007, Elter et al. [14] conducted a study of CAD to see its effectiveness in reducing unnecessary biopsies. Using two databases, 2,100 images from the United States and 961 images from the United Kingdom, which was collected in the year 2003-2006. It was found to help doctors decide to reduce unnecessary biopsy. However, the study used databases from the United States and the United Kingdom and information from many decades ago. Therefore, we were interested in collecting such information in Thailand and contemporary data.

This research is a study of a data set from Phayao Hospital, northern Thailand. The work has done by collecting BI-RADS data on patient age, mass shape, margin, and density of the mass including the pathology, for example see in Figures 4 and 5. Then matching the results as benign or malignant correlated with pathological results. Then enter these information to train machine learning using a new algorithm to find accurate prediction models from mammography images in order to further develop a prediction prototype using real data from Thailand.

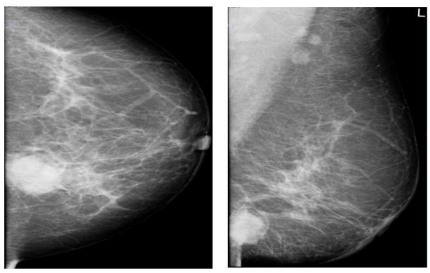

Figure 4: Example of mammography; left craniocaudal (CC) and mediolateral oblique (MLO) views of a 51-year-old woman, shows a lobulate-shaped, spiculate high density mass at left inner-lower region. The BI-RADS assessment = 4c, Shape = 3, Margin = 5, Density =1, and Severity =1. The pathological proven to be invasive ductal carcinoma.

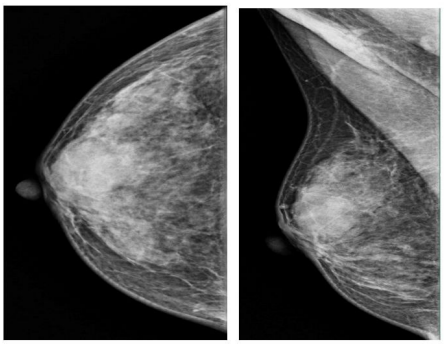

Figure 5: Example of mammography; right craniocaudal (CC) and mediolateral oblique (MLO) views of a 41-year-old woman, shows a round-shaped, circumscribed iso density mass at right upper-central region. The BI-RADS assessment = 2, Shape = 1, Margin = 1, Density =2, and Severity = 0. The ultrasonography proven to be a cyst.

BI-RADS can be applied to mammography, ultrasound, and MRI. This article reflects the 5th edition, published in 2013 [31]. Breast imaging studies are assigned one of the following seven assessment categories, see in Table 4.

BI-RADS	
6	Known biopsy-proven malignancy
5	Highly suggestive of malignancy >95% probability of malignancy appropriate action should be taken
4	Suspicious for malignancy 2-95% probability of malignancy for mammography and ultrasound, these can be further divided: 4A: low suspicion for malignancy (2-9%) 4B: moderate suspicion for malignancy (10-49%) 4C: high suspicion for malignancy (50-95
3	probably benign <2% probability of malignancy short interval follow-up suggested
2	Benign 0% probability of malignancy
1	Negative symmetrical and no masses, architectural distortion, or suspicious calcifications
0	Incomplete need additional imaging evaluation (additional mammographic views or ultrasound) and/or for mammography, Obtaining previous images not available at the time of reading

Table 4: BI-RADS of assessment categories of breast cancer.

To obtain the efficiency model of BIRADS prediction, mammographic masses dataset from UCI machine learning repository [38] are used with dataset from Phayao Hospital into three cases. The overview of three cases are shown as in Table 5.

Table 5: Data overview of three cases.

	Instances		BI-RADS	Age	Shape	Margin	Density	Severity
Phayao hospital (D1)	171	Mean	4.40	54.42	2.78	3.22	1.22	0.67
		SD	0.93	11.90	1.27	1.69	0.44	0.47
		Max	6	85	4	5	3	1
		Min	2	10	1	1	1	0
Phayao hospital and UCI dataset* (D2)	176	Mean	4.40	54.66	2.80	3.24	1.27	0.67
		SD	1.17	11.97	1.27	1.68	0.53	0.47
		Max	6	85	4	5	3	1
		Min	0	10	1	1	1	0
UCI dataset after deleting data missing (D3)	828	Mean	4.33	55.81	2.78	2.82	2.92	0.48
		SD	0.69	14.68	1.24	1.57	0.35	0.50
		Max	6	96	4	5	4	4
		Min	0	18	1	1	1	0

*Only five data of BI-RADS 0 from the UCI dataset are added since the Phayao Hospital dataset is without BI-RADS 0 and BI-RADS 1.

In this section, we shall apply Algorithms 3.4 and 3.5 in the process of machine learning data classification by using an extreme learning machine (ELM) [19], which was defined as follows: let $u := \{(a^k, t^k) : a^k \in R^m, t^k \in R^p, k = 1, 2, \ldots, M\}$ be a set of training data, where M is distinct samples, a^k and t^k are an input training data and training target target, respectively. The output function of ELM with sigmoid activation function for a standard single hidden layer feedforward networks (SLFNs) by N hidden nodes are mathematically modeled as:

$$O^j = \sum_{i=1}^{N} \omega^i \frac{1}{1 + e^{-(w^i a^j + b^i)}},$$

where w^i, b^i are parameter weight and bias random inputting, respectively, and ω^i is the optimal output weight. To find the vector ω^i at the i-th hidden node. The above equations can be denoted in the form of matrix as

$$A = \begin{bmatrix} \frac{1}{1+e^{-(w^1 a^1 + b^1)}} & \cdots & \frac{1}{1+e^{-(w^N a^1 + b^N)}} \\ \vdots & \ddots & \vdots \\ \frac{1}{1+e^{-(w^1 a^M + b^1)}} & \cdots & \frac{1}{1+e^{-(w^N a^M + b^N)}} \end{bmatrix},$$

A is called the hidden layer output matrix, $T = [t^{1^T}, \ldots, t^{N^T}]^T$ represents the training target data matrix, and $\omega = [\omega^{1^T}, \ldots, \omega^{N^T}]^T$ represents optimal output weight such that the linear system $A\omega = T$, where $A^\dagger$. In the case of $A\omega = T$ does not have a solution, the best approximate solution is considered which is called least squares problem.

Regularization of least square problems is a technique commonly used in machine learning to avoid overfitting of models, which can lead to better predict and suitable solution in classification problems. We shall consider the regularization of least square problems in three models, including viewing the constrained closed convex subset C with setting the problems (3.6) to obtain these three goals.

(i) Regularization of least squares problem by L_1 (RL-L_1) or well-known called the least absolute shrinkage and selection operator (LASSO): for $\lambda > 0$,

$$\min_{\omega \in R^N} \frac{1}{2}\|A\omega - T\|_2^2 + \lambda\|\omega\|_1. \tag{4.1}$$

Set $F(\omega) \equiv \nabla(\frac{1}{2}\|A\omega - T\|_2^2)$, $G(\omega) \equiv \partial(\lambda\|\omega\|_1)$, $C = H$.

(ii) Regularization of least square problem by L_2 (RL-L_2): for $\lambda > 0$,

$$\min_{\omega \in R^N} \frac{1}{2}\|A\omega - T\|_2^2 + \lambda\|\omega\|_2^2. \tag{4.2}$$

Set $F(\omega) \equiv \nabla(\frac{1}{2}\|A\omega - T\|_2^2)$, $G(\omega) \equiv \partial(\lambda\|\omega\|_2^2)$, $C = H$.

(iii) Regularization of least square problem by L_1 with constrained closed convex set by L_1 (RLC-L_1): for $\lambda, \rho > 0$,

$$\min_{\omega \in C} \frac{1}{2}\|A\omega - T\|_2^2 + \lambda\|\omega\|_1, \tag{4.3}$$

where $C = \{\omega : \|\omega\|_1 \leqslant \rho\}$. Set $F(\omega) \equiv \nabla(\frac{1}{2}\|A\omega - T\|_2^2)$, $G(\omega) \equiv \partial(\lambda\|\omega\|_1)$, $C = \{\omega : \|\omega\|_1 \leqslant \rho\}$.

We examined four assessment criteria, including accuracy, precision, recall, and F1-score [47], to assess the effectiveness of the classification algorithms. These metrics are defined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \times 100\%, \qquad \text{Precision} = \frac{TP}{TP + FP} \times 100\%,$$
$$\text{Recall} = \frac{TP}{TN + FN} \times 100\%, \qquad F1 - \text{score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}.$$

In these matrices, TP is the True Positive, TN is the True Negative, FP is the False Positive, and FN is the False Negative.

The multi-class cross-entropy loss [7] is a metric used in classification tasks to evaluate a model's ability to distinguish between multiple classes. This measurement is determined by calculating the following average:

$$\text{Loss} = -\sum_{i=1}^{N} \varphi_i \log \bar{\varphi}_i,$$

where $\bar{\varphi}_i$ denotes the i-th scalar value in the model's output, while φ_i signifies the corresponding target value associated with that particular scalar. The variable N represents the total count of scalar values encompassed within the entire model's output.

Next, three cases of the dataset are split into 80% for training and 20% for validation. Input weight and bias are randomized in the interval [-0.5,0.5] and [-0.1,0.1], respectively, with 150 hidden nodes. To obtain the highest performance of the our algorithm, we shall consider the necessary parameters of Algorithms 3.4 and 3.5 for (4.1)-(4.3) models as in Table 6 when θ^k and δ^k are the same as in (3.7) and (3.8), respectively.

From Tables 7-9, the highest accuracy, 88.57 %, is obtained when the Phayao Hospital and UCI dataset (D2) is used, with F1-score, Recall, and Precision for all Algorithm 3.4, Algorithm 3.5, and FISTA algorithm. Our Algorithm 3.4 and Algorithm 3.5 by (4.1)-(4.3) models achieved less of number of iterations than FISTA algorithm, and by (4.2)-(4.3) models also achieved less of training time than FISTA algorithm.

Table 6: The necessary parameters of Algorithms 3.4 and 3.5 for (4.1)-(4.3) models.

	Model	γ	α^k	β^k	θ	δ	λ	ρ
Algorithm 3.4	RL-L_1	$\frac{1.999}{\max(\text{eigenvalue}(A^TA))}$	$\frac{k}{2k+1}$	$\frac{k}{1.5k+1}$	θ_{FISTA}	-0.001	0.1	-
	RL-L_2	$\frac{1.999}{\max((\text{eigenvalue}(A^TA))}$	$\frac{k}{2k+1}$	$\frac{k}{1.5k+1}$	θ_{FISTA}	-0.001	0.1	-
	RLC-L_1	$\frac{1.999}{\max((\text{eigenvalue}(A^TA))}$	$\frac{k}{2k+1}$	$\frac{k}{1.5k+1}$	θ_{FISTA}	-0.001	0.1	10
Algorithm 3.5	RL-L_1	$\frac{1.999}{\max((\text{eigenvalue}(A^TA))}$	$\frac{k}{2k+1}$	$\frac{k}{2k+1}$	θ_{FISTA}	-0.001	0.1	-
	RL-L_2	$\frac{1.999}{\max((\text{eigenvalue}(A^TA))}$	$\frac{k}{2k+1}$	$\frac{k}{2k+1}$	θ_{FISTA}	-0.001	0.1	-
	RLC-L_1	$\frac{1.999}{\max((\text{eigenvalue}(A^TA))}$	$\frac{k}{2k+1}$	$\frac{k}{2k+1}$	θ_{FISTA}	-0.001	0.1	8
FISTA	RL-L_1	$\frac{1}{\|A\|^2}$	-	-	-	-	0.001	-
	RL-L_2	$\frac{1}{\|A\|^2}$	-	-	-	-	0.001	-

Table 7: The assessment results of Algorithms 3.4 and 3.5 comparing with FISTA algorithm by using Phayao Hospital dataset (D1).

	Model	Number of iterations	Training time	Accuracy	F1-score	Recall	Precision
Algorithm 3.4	RL-L_1	305	0.0363	87.06	90.00	88.99	91.04
	RL-L_2	262	0.0498	87.06	90.00	88.99	91.04
	RLC-L_1	305	0.0361	87.06	90.00	88.99	91.04
Algorithm 3.5	RL-L_1	305	0.0351	87.06	90.00	88.99	91.04
	RL-L_2	262	0.0313	87.06	90.00	88.99	91.04
	RLC-L_1	305	0.0440	87.06	90.00	88.99	91.04
FISTA	RL-L_1	366	0.0403	87.06	90.00	88.99	91.04
	RL-L_2	366	0.0426	87.06	90.00	88.99	91.04

Table 8: The assessment results of Algorithm 3.4 and Algorithm 3.5 compare with FISTA algorithm by using Phayao Hospital and UCI dataset (D2).

	Model	Number of iterations	Training Time	Accuracy	F1-score	Recall	Precision
Algorithm 3.4	RL-L_1	293	0.0623	88.57	91.37	90.54	92.22
	RL-L_2	258	0.0365	88.57	91.37	90.54	92.22
	RLC-L_1	293	0.0383	88.57	91.37	90.54	92.22
Algorithm 3.5	RL-L_1	293	0.0603	88.57	91.37	90.54	92.22
	RL-L_2	258	0.0337	88.57	91.37	90.54	92.22
	RLC-L_1	293	0.0355	88.57	91.37	90.54	92.22
FISTA	RL-L_1	361	0.0416	88.57	91.37	90.54	92.22
	RL-L_2	361	0.0431	88.57	91.37	90.54	92.22

Table 9: The assessment results of Algorithm 3.4 and Algorithm 3.5 compare with FISTA algorithm by using UCI dataset after deleting data missing (D3).

	Model	Number of iterations	Training Time	Accuracy	F1-score	Recall	Precision
Algorithm 3.4	RL-L_1	79	0.0385	84.43	86.21	86.00	86.42
	RL-L_2	79	0.0389	84.43	86.21	86.00	86.42
	RLC-L_1	79	0.0381	84.43	86.21	86.00	86.42
Algorithm 3.5	RL-L_1	79	0.0413	84.43	86.21	86.00	86.42
	RL-L_2	79	0.0429	84.43	86.21	86.00	86.42
	RLC-L_1	79	0.0430	84.43	86.21	86.00	86.42
FISTA	RL-L_1	82	0.0418	84.63	79.15	82.69	75.89
	RL-L_2	82	0.0430	84.63	79.15	82.69	75.89

Based on the conclusion of Tables 7-9, we shall show the accuracy and loss plots of our Algorithm 3.4 and Algorithm 3.5 by (4.1)-(4.3) models as in Figures 6-17.

From Figures 6-17, we see that our Algorithms 3.4 and 3.5 by (4.1)-(4.3) models are well fitting model for the Phayao Hospital and UCI dataset (D2).

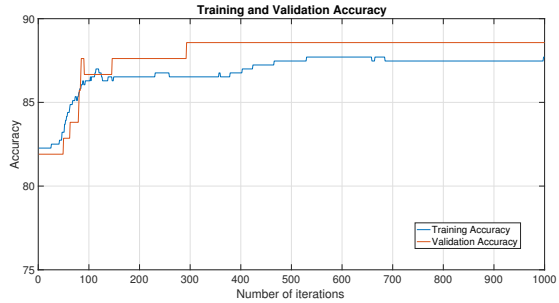

Figure 6: The accuracy plots of Algorithm 3.4 by RL-L_1 (4.1) model.

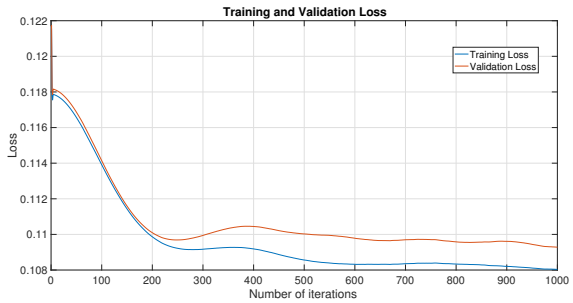

Figure 7: The loss plots of Algorithm 3.4 by RL-L_1 (4.1) model.

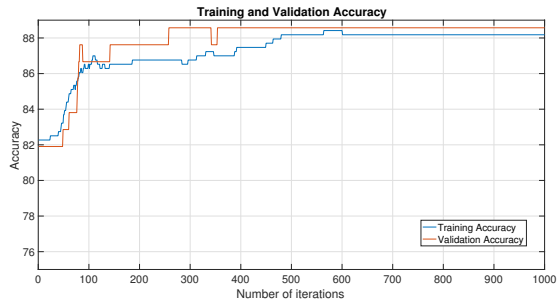

Figure 8: The accuracy plots of Algorithm 3.4 by RL-L_2 (4.2) model.

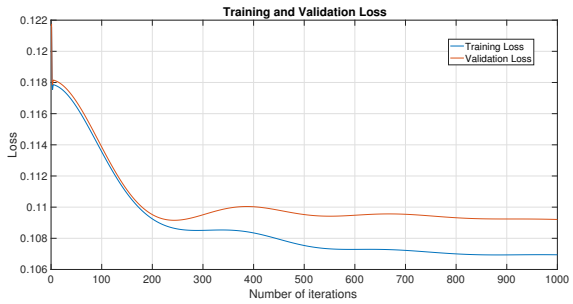

Figure 9: The loss plots of Algorithm 3.4 by RL-L_2 (4.2) model.

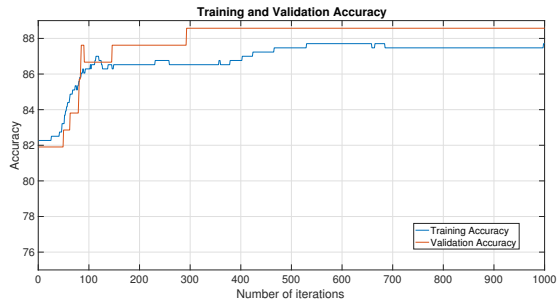

Figure 10: The accuracy plots of Algorithm 3.4 by RLC-L_1 (4.3) model.

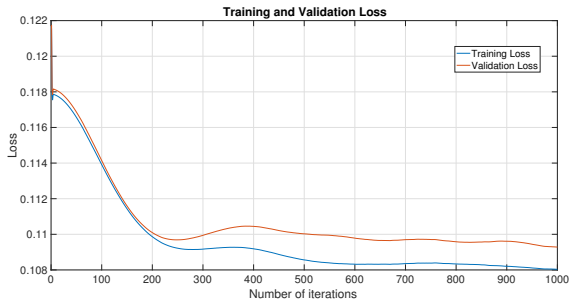

Figure 11: The loss plots of Algorithm 3.4 by RLC-L_1 (4.3) model.

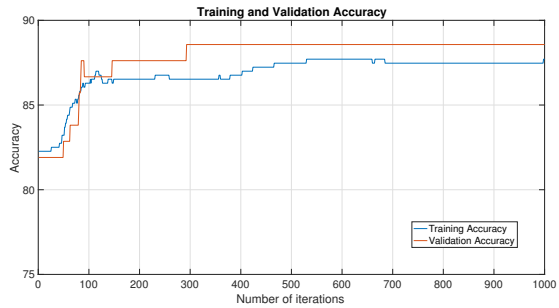

Figure 12: The accuracy plots of Algorithm 3.5 by RL-L_1 (4.1) model.

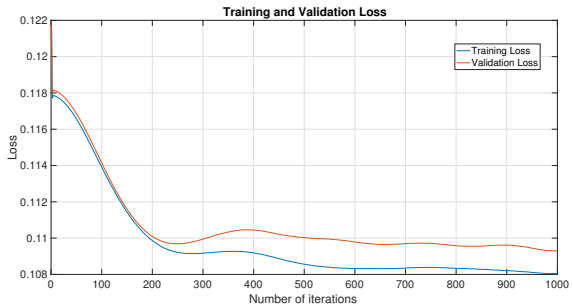

Figure 13: The loss plots of Algorithm 3.5 by RL-L_1 (4.1) model.

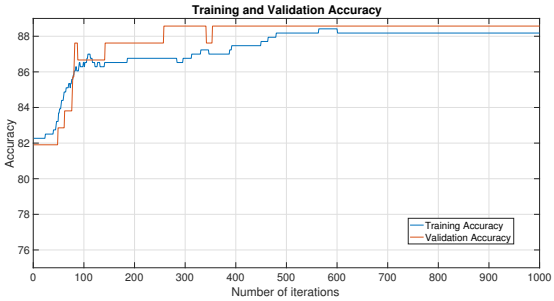

Figure 14: The accuracy plots of Algorithm 3.5 by RL-L_2 (4.2) model.

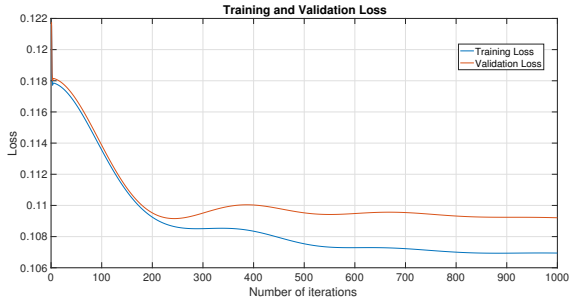

Figure 15: The loss plots of Algorithm 3.5 by RL-L_2 (4.2) model.

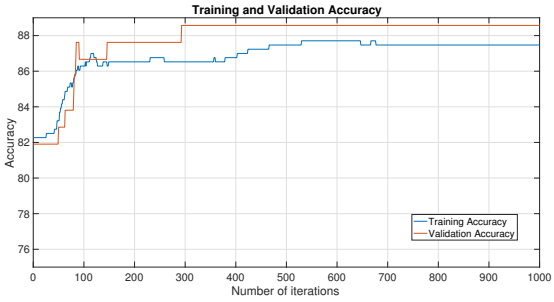

Figure 16: The accuracy plots of Algorithm 3.5 by RLC-L_1 (4.3) model.

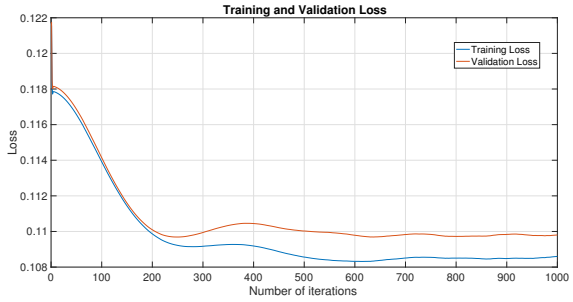

Figure 17: The loss plots of Algorithm 3.5 by RLC-L_1 (4.3) model.

5. Conclusions

This paper presents a novel double inertial embedded Ishikawa algorithm designed to determine a common fixed point for two distinct nonexpansive mappings within a real Hilbert space. We establish a weak convergence theorem based on appropriate conditions and demonstrate the utility of our approach using a finite-dimensional space to support our primary findings. Our algorithm is particularly useful for addressing variational inclusion problems by leveraging newly developed double inertial projection forward-backward splitting algorithms, adaptable to constrained closed convex subsets.

Moreover, we apply these algorithms to determine the optimal output weight for the Extreme Learning Machine (ELM) model, using a real dataset from Phayao Hospital in Northern Thailand. The empirical results are promising: our algorithm achieved up to 88.57% accuracy, 92.22% precision, 90.54% recall, and 91.37% F1-score. This performance was obtained with only 171 real data samples from Phayao Hospital and 5 BI-RADS 0 data samples from the UCI website, requiring less than one minute of training time. This highlights not only the algorithm's efficiency but also its potential to significantly benefit CPU utilization. Expanding the data collection scope could lead to the development of more sophisticated models applicable on a larger scale.

Data availability

(i) The real data set from Phayao Hospital are available from the corresponding author (W. Cholamjiak), upon reasonable request.

(ii) The mammographic mass dataset from UCI is available on the UCI website (http://archive.ics.uci.edu/ml/datasets/mammographic+mass).

Institutional review board statement

This study was conducted in accordance with the Declaration of Helsinki, the Belmont report, CIOMS Guideline international conference on Harmonization in Good Cinical Pratice or ICH-GCP and with

approval from the Ethics Committee and Institutional Review Board of Phayao Hospital (Institutional Review Board (IRB) approval, IRB Number: COA No.20966-02-26).

Funding

National Research Council of Thailand and University of Phayao (N42A650334), and University of Phayao and Thailand Science Research and Innovation Fund (Fundamental Fund 2025, Grant No. 5013 /2567).

Acknowledgments

The authors would like to thank the Phayao Hospital and the University of Phayao.

References

[1] J. Abounadi, D. P. Bertsekas, V. Borkar, *Stochastic approximation for nonexpansive maps: application to Q-learning algorithms*, SIAM J. Control Optim., **41** (2002), 1–22. 1
[2] H. Attouch, A. Cabot, *Convergence rates of inertial forward-backward algorithms*, SIAM J. Optim., **28** (2018), 849–874. 1
[3] T. Ayer, M. U. Ayvaci, Z. X. Liu, O. Alagoz, and E. S. Burnside, *Computer-aided diagnostic models in breast cancer screening*, Imaging Med., **2** (2010), 313–323. 4
[4] H. H. Bauschke, J. M. Borwein, *On projection algorithms for solving convex feasibility problems*, SIAM Rev., **38** (1996), 367–426. 1
[5] H. H. Bauschke, P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, Springer, New York, (2011). 1
[6] A. Beck, M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAM J. Imaging Sci., **2** (2009), 183–202. 1
[7] M. Ben naceur, M. Akil, R. Saouli, R. Kachouri, *Fully automatic brain tumor segmentation with deep learning-based selective attention using overlapping patches and multi-class weighted cross-entropy*, Med. Image Anal., **63** (2020). 4
[8] R. I. Boţ, E. R. Csetnek, S. C. László, *An inertial forward–backward algorithm for the minimization of the sum of two nonconvex functions*, EURO J. Comput. Optim., **4** (2016), 3–25. 1
[9] F. E. Browder, *Fixed-point theorems for noncompact mappings in Hilbert space*, Proc. Nat. Acad. Sci. U.S.A., **53** (1965), 1272–1276. 1
[10] L. C. Ceng, P. Cubiotti, J. C. Yao, *Strong convergence theorems for finitely many nonexpansive mappings and applications*, Nonlinear Anal., **67** (2007), 1464–1473. 1
[11] Y.-W. Chang, J. K. An, N. Choi, K. H. Ko, K. H. Kim, K. Han, J. K. Ryu, *Artificial intelligence for breast cancer screening in mammography (AI-STREAM): A prospective multicenter study design in Korea using AI-based CADe/x*, J. Breast Cancer, **25** (2022), 57–68. 4
[12] A. Columbu, R. Díaz Fuentes, S. Frassu, *Uniform-in-time boundedness in a class of local and nonlocal nonlinear attraction–repulsion chemotaxis models with logistics*, Nonlinear Anal. Real World Appl., **79** (2024), 14 pages. 1
[13] G. Durhan, A. Azizova, Ö. Önder, K. Kösemehmetoğlu, J. Karakaya, M. G. Akpınar, F. Demirkazık, A. Üner, *Imaging findings and clinicopathological correlation of breast cancer in women under 40 years old*, Eur. J. Breast Health, **15** (2019), 147–152. 4
[14] M. Elter, R. Schulz-Wendtland, T. Wittenberg, *The prediction of breast cancer biopsy outcomes using two CAD approaches that both emphasize an intelligible decision process*, Med. Phys., **34** (2007), 4164–4172. 4
[15] K. Freeman, J. Geppert, C. Stinton, D. Todkill, S. Johnson, A. Clarke, S. Taylor-Phillips, *Use of artificial intelligence for image analysis in breast cancer screening programmes: Systematic review of test accuracy*, BMJ, **374** (2021), 15 pages. 4
[16] K. Goebel, W. A. Kirk, *Topics in metric fixed point theory*, Cambridge University Press, Cambridge, (1990). 1, 2.1
[17] L. J. Grimm, C. S. Avery, E. Hendrick, J. A. Baker, *Benefits and risks of mammography screening in women ages 40 to 49 years*, J. Prim. Care Community Health, **13** (2022), 1–6. 4
[18] Z. Guo, J. Xie, Y. Wan, M. Zhang, L. Qiao, J. Yu, S. Chen, B. Li, Y. Yao, *A review of the current state of the computer-aided diagnosis (CAD) systems for breast cancer diagnosis*, Open Life Sci., **17** (2022), 1600–1611. 4
[19] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, *Extreme learning machine: Theory and applications*, Neurocomputing, **70** (2006), 489–501. 4
[20] S. Ishikawa, *Fixed points by a new iteration method*, Proc. Amer. Math. Soc., **44** (1974), 147–150. 1
[21] Y. Jiang, *Computer-aided diagnosis of breast cancer in mammography: evidence and potential*, Technol. Cancer Res. Treat., **1** (2002), 211–216. 4

[22] Y. Jiang, R. M. Nishikawa, R. A. Schmidt, C. E. Metz, M. L. Giger,, K. Doi, *Improving breast cancer diagnosis with computer-aided diagnosis*, Acad. Radiol., **6** (1999), 22–33. 4

[23] Z. Jiao, I. Jadlovská, T. Li, *Finite-time blow-up and boundedness in a quasilinear attraction–repulsion chemotaxis system with nonlinear signal productions*, Nonlinear Anal. Real World Appl., **77** (2024), 15 pages. 1

[24] Z. Jiao, I. Jadlovská, T. Li, *Global existence in a fully parabolic attraction-repulsion chemotaxis system with singular sensitivities and proliferation*, J. Differ. Equ., **411** (2024), 227–267. 1

[25] J. Katzen, K. Dodelzon, *A review of computer aided detection in mammography*, Clin. Imaging, **52** (2018), 305–309. 4

[26] W. A. Kirk, *A fixed point theorem for mappings which do not increase distances*, Amer. Math. Monthly, **72** (1965), 1004–1006. 1

[27] M. Larsen, C. F. Aglen, S. R. Hoff, H. Lund-Hanssen, S. Hofvind, *Possible strategies for use of artificial intelligence in screen-reading of mammograms, based on retrospective data from 122,969 screening examinations*, Eur. Radiol., **32** (2022), 8238–8246. 4

[28] T. Li, S. Frassu, G. Viglialoro, *Combining effects ensuring boundedness in an attraction–repulsion chemotaxis model with production and consumption*, Z. Angew. Math. Phys., **74** (2023), 21 pages. 1

[29] T. Li, G. Viglialoro, *Boundedness for a nonlocal reaction chemotaxis model even in the attraction-dominated regime*, Differ. Integral Equ., **34** (2021), 315–336. 1

[30] J. Liang, C. B. Schönlieb, *Improving FISTA: Faster, Smarter, and Greedier*, arXiv preprint arXiv:1811.01430, **2018** (2018). 1

[31] L. Liberman, J. H. Menell, *Breast imaging reporting and data system (BI-RADS)*, Radiol. Clin., **40** (2002), 409–430. 4

[32] K. Loizidou, R. Elia, C. Pitris, *Computer-aided breast cancer detection and classification in mammography: A comprehensive review*, Comput. Biol. Med., **153** (2023), 24 pages. 4

[33] G. López, V. Martín-Márquez, F. Wang, H.-K. Xu, *Forward-backward splitting methods for accretive operators in Banach spaces*, Abstr. Appl. Anal., **2012** (2012), 25 pages. i

[34] D. A. Lorenz, T. Pock, *An inertial forward-backward algorithm for monotone inclusions*, J. Math. Imaging Vision, **51** (2015), 311–325. 1

[35] S. Lukasiewicz, M. Czeczelewski, A. Forma, J. Baj, R. Sitarz, A. Stanisławek, *Breast cancer—Epidemiology, risk factors, classification, prognostic markers, and current treatment strategies—An updated review*, Cancers, **13** (2021), 30 pages. 4

[36] W. R. Mann, *Mean value methods in iteration*, Proc. Amer. Math. Soc., **4** (1953), 506–510. 1

[37] R. M. Mann, R. Hooley, R. G. Barr, L. Moy, *Novel approaches to screening for breast cancer*, Radiology, **297** (2020), 266–285. 4

[38] E. Matthias, *Mammographic Mass*, UCI Machine Learning Repository, (2007). 4

[39] P. Nabheerong, W. Kiththiworaphongkich, W. Cholamjiak, *Breast Cancer Screening Using a Modified Inertial Projective Algorithm for Split Feasibility Problems*, Int. J. Breast Cancer, **2023** (2023), 9 pages. 1

[40] E. U. Ofoedu, *Strong convergence theorem for uniformly* L*-Lipschitzian asymptotically pseudocontractive mapping in real Banach space*, J. Math. Anal. Appl., **321** (2006), 722–728. 2.3

[41] Z. Opial, *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bull. Amer. Math. Soc., **73** (1967), 591–597. 2.2

[42] P. Peeyada, R. Suparatulatorn, W. Cholamjiak, *An inertial Mann forward-backward splitting algorithm of variational inclusion problems and its applications*, Chaos Solitons Fractals, **158** (2022), 7 pages. 1

[43] B. T. Polyak, *Some methods of speeding up the convergence of iteration methods*, USSR Comput. Math. Math. Phys., **4** (1964), 1–17. 1

[44] X. Qin, S. Y. Cho, H. Zhou, *Common fixed points of a pair of non-expansive mappings with applications to convex feasibility problems*, Glasg. Math. J., **52** (2010), 241–252. 1

[45] J. Shiraishi, Q. Li, D. Appelbaum, K. Doi, *Computer-aided diagnosis and artificial intelligence in clinical imaging*, Semin. Nucl. Med., **41** (2011), 449–462. 4

[46] N. X. Tan, *On the existence of solutions of quasivariational inclusion problems*, J. Optim. Theory Appl., **123** (2004), 619–638.

[47] N. W. S. Wardhani, M. Y. Rochayani, A. Iriany, A. D. Sulistyono, P. Lestantyo, *Cross-validation metrics for evaluating classification performance on imbalanced data*, In: 2019 International Conference on Computer, Control, Informatics and its Applications (IC3INA), IEEE, (2019), 14–18. 4

[48] S.-S. Zhang, J. H. Lee, C. K. Chan, *Algorithms of common solutions to quasi variational inclusion and fixed point problems*, Appl. Math. Mech. (English Ed.), **29** (2008), 571–581. 1