Article

Matrix Expression of Convolution and Its Generalized Continuous Form

Young Hee Geum [1,†]**, Arjun Kumar Rathie** [2,†] **and Hwajoon Kim** [3,*,†]

1   Department of Applied Mathematics, Dankook University, Cheonan 31116, Korea; conpana@empal.com
2   Department of Mathematics, Vedant College of Engineering & Technology (Rajasthan Technical University), Kota 324010, India; arjunkumarrathie@gmail.com
3   Department of IT Engineering, Kyungdong University, Yangju 11458, Korea
*   Correspondence: math@kduniv.ac.kr or cellmath@gmail.com
†   These authors contributed equally to this work.

check for updates

Abstract: In this paper, we consider the matrix expression of convolution, and its generalized continuous form. The matrix expression of convolution is effectively applied in convolutional neural networks, and in this study, we correlate the concept of convolution in mathematics to that in convolutional neural network. Of course, convolution is a main process of deep learning, the learning method of deep neural networks, as a core technology. In addition to this, the generalized continuous form of convolution has been expressed as a new variant of Laplace-type transform that, encompasses almost all existing integral transforms. Finally, we would, in this paper, like to describe the theoretical contents as detailed as possible so that the paper may be self-contained.

Keywords: matrix expression of convolution; Laplace-type transforms; convolution neural network; kernel

1. Introduction

Deep learning means the learning of deep neural networks, called deep and if multiple hidden layers exist. Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction [1]. The convolution in convolutional deep neural network (CNN) is the tool for obtaining a feature map from the original image data, it sweeps the original image with a kernel matrix, and transforms the original data into a different shape. This distorted image is called a feature map. Therefore, in CNN, the convolution can be regarded as a tool that creates a feature map from the original image. Herein, the concept of convolution in artificial intelligence is demonstrated mathematically.

The core concept of CNN is the convolution which applies the weight to the receptive fields only, and it transforms the original data into a feature map. This process is called convolution. This is a similar principle to integral transform. The method of Integral transform maps from the original domain to another domain to solve a given problem more easily. Since the matrix expression of convolution is an essential concept in artificial intelligence, we believe that this study would certainly be meaningful. In addition to this, the generalized continuous form of convolution has also been studied, and thus this form is expressed as a new variant of Laplace-type transform.

On one hand, the transform theory is extensively utilized in fields involving medical diagnostic equipment, such as magnetic resonance imaging or computed tomography. Typically, a projection data are obtained by an integral transform, and an image using an inverse transform is produced. Although plausible integral transforms exist, almost all existing integral transforms are not sufficiently satisfied with fullness, and can be interpreted as a Laplace-type

transform. One of us proposed a comprehensive form of the Laplace-type integral transform in [2]. The present study is being conducted to investigate the matrix expression of convolution and its generalized continuous form.

In [2], a Laplace-type integral transform was proposed, expressed as

$$G_\alpha(f) = G(f) = u^\alpha \int_0^\infty e^{-\frac{t}{u}} f(t) dt. \tag{1}$$

For values of α as 0, -1, 1, and -2, we have, respectively, the Laplace [3], Sumudu [4], Elzaki [5], and Mohand transforms [6]. This form can be expressed in various manners. Replacing t by ut, we have

$$G(f) = u^\beta \int_0^\infty e^{-t} f(ut) dt,$$

where $\beta = \alpha + 1$. In the form, β values of 1, 0, 2, and -1 correspond to the Laplace, Sumudu, Elzaki, and Mohand transforms, respectively. If we substitute $u = 1/s$ in (1), we then obtain the simplest form of the generalized integral transform as follows:

$$G(f) = s^\gamma \int_0^\infty e^{-st} f(t) dt,$$

where $\gamma = -\alpha$. In this form, the Laplace, Sumudu, Elzaki, and Mohand transforms have γ values of 0, 1, -1, and 2, respectively. It is somewhat paved, but essentially a simple way to derive the Sumudu transform is to multiply the Laplace transform by s. Similarly, it can be obtained multiply by s^{-1} to obtain the Elzaki transform, and multiply by s^2 to obtain the Mohand transform. The natural transform [7] can be obtained by substituting $f(t)$ with $f(ut)$. Additionally, by substituting $t = \ln x$, the Laplace-type transform $G(f)$ can be expressed as

$$s^{-\alpha} \int_1^\infty f(\ln x) x^{-s-1} \, dx.$$

As a similar form, there is a Mellin transform [8] of the form

$$\int_0^\infty f(x) x^{s-1} \, dx.$$

As shown above, many integral transforms have their own fancy masks, but most of them can essentially be interpreted as Laplace-type transforms. From a different point of view, a slight change in the kernel results in a significant difference in the integral transform theory. Meanwhile, plausible transforms exist, such as the Fourier, Radon, and Mellin transforms. Typically, if the interval of integration and the power of kernel are different, it can be interpreted as a completely different transform. Studies using Laplace transform were conducted in [9,10]. The generalized solutions of the third-order Cauchy–Euler equation in the space of right-sided distributions has found [9], studied the solution of the heat equation without boundary conditions [10], and investigated further properties of Laplace-type transform [11]. As an application, a new class of Laplace-type integrals involving generalized hypergeometric functions has been studied [12,13]. As for research related to the integral equation, Noeiaghdam et al. [14] presented a new scheme based on the stochastic arithmetic. The scheme is presented to guarantee the validity and accuracy of the homotopy analysis method. Different kinds of integral equations such as singular and first kind are considered to find the optimal results by applying the proposed algorithms.

The main objective of this study is to investigate the matrix expression of convolution and its generalized continuous form. The generalized continuous form of the matrix expression was carried out in the form of a new variant of Laplace-type transform. The obtained result are as follows:

(1) If the matrix representing the function (image) f is A and the matrix representing the function g is B, then the convolution $f * g$ is represented by *the sum of all elements of $A \circ B$* and this is the same as $tr(AB^T)$ where $\circ$ is array multiplication, T is the transpose, and tr is the trace. Thus, the convolution in artificial intelligence (AI) is the same as $tr(AB^T)$.

(2) The generalized continuous form of the convolution in AI can be represented as

$$V(f) = \Phi(u) \int_0^\infty e^{-t\Delta} f(t) \, dt,$$

where $\Phi(u)$ is an arbitrary bounded function and

$$\Delta = \Delta(\delta, u) = \frac{\ln[1 + \frac{\delta - 1}{u}]}{\delta - 1}.$$

2. Matrix Expression of Convolution in Convolutional Neural Network (CNN)

Note that functions can be interpreted as images in artificial intelligence (AI). The convolution is changed from $\int_0^t f(\tau)g(t - \tau)d\tau$ to

$$\sum_{\tau=0}^t f(\tau)g(t - \tau)$$

by the discretization. The convolution in CNN is the tool for obtaining a feature map from the original image data, plays a role to sweeping the original image with kernel matrices (or filter), and it transforms original data into a different shape. In order to calculate the convolution, each $n \times n$ part of the original matrix is element-wise multiplied by the kernel matrix and all its components are added. Typically, the kernel matrix is using by 3×3 matrix. On the one hand, the pooling (or sub-sampling) is a simple job, reducing the size of the image made by convolution. It is the principle that the resolution is increased when the screen is reduced.

Let the matrix representing the function f is A and the matrix representing the function g is B. For two matrices A and B of the same dimension, the array multiplication (or sweeping) $A \circ B$ is given by

$$(A \circ B)_{ij} = (A)_{ij}(B)_{ij}.$$

For example, the array multiplication for 2×2 matrices is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \circ \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae & bf \\ cg & dh \end{pmatrix}.$$

The array multiplication appears in lossy compression such as joint photographic experts group and the decoding step. Let us look at an example.

Example 1. *In the classification field of AI, the pixel is treated as a matrix. When the original image is*

$$\begin{pmatrix} 1 & 2 & 3 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 3 & 0 & 1 & 0 & 1 \\ 1 & 0 & 2 & 1 & 0 \\ 0 & 1 & 1 & 2 & 0 \end{pmatrix},$$

array multiplying the kernel matrix

$$\begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix}$$

on the first 3 × 3 matrix, we obtain the matrix

$$\begin{pmatrix} 2 & 0 & 3 \\ 0 & 1 & 4 \\ 6 & 0 & 1 \end{pmatrix}.$$

Now, adding all of its components, we obtain 17. *Next, if we array-multiply the kernel matrix to the 3 × 3 matrix*

$$\begin{pmatrix} 2 & 3 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

on the right and add all the components, we get 8 *by stride* 1. *If we continue this process to the final matrix*

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 0 \end{pmatrix},$$

we get 6. *Consequently, the original matrix changes to*

$$\begin{pmatrix} 17 & 8 & 10 \\ 8 & 5 & 10 \\ 12 & 8 & 6 \end{pmatrix}$$

by using the convolution kernel. This is called the convolved feature map.

This is just an example for understanding, and in perceptron the output uses a value between -1 and 1 using the activation function. Note that the perceptron is an artificial network designed to mimic the brain's cognitive abilities. Therefore, the output of neuron (or node) Y can be represented as

$$Y = sign\left[\sum_{i=1}^{n} x_i w_i - \Theta \right] = \begin{cases} 1 & X \geq \Theta \\ -1 & X < \Theta \end{cases},$$

where w is a weight, Θ is the threshold value, and X is the activation function with $X = \sum_{i=1}^{n} x_i w_i$. In the backpropagation algorithm of deep neural network, the sigmoid function

$$Y^{sigmoid} = \frac{1}{1 + e^{-x}}$$

is used as the activation function [15]. This function is easy to differentiate and ensures neuron output is in $[0,1]$. If max-pulling is applied to the above convolved feature map, the resulting matrix becomes 1×1 matrix $(17) = 17$.

As discussed above, convolution in AI can be obtained by array multiplication. We would like to associate this definition with matrix multiplication in mathematics.

Definition 1. *(Convolution in AI) If the matrix representing the function (image) f is A and the matrix representing the function g is B, then the convolution f ∗ g is represented by the sum of all elements of A ∘ B and this is the same as tr(AB^T) where ∘ is array multiplication, T is the transpose, and tr is the trace. Thus, the convolution in AI is the same as tr(AB^T), the sum of all elements on the diagonal with the right side facing down in AB^T.*

Typically, the convolution kernel is used as a 3×3 matrix, but for easy understanding, let us consider a 2×2 matrix.

Example 2. *If*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

and

$$B = \begin{pmatrix} e & f \\ g & h \end{pmatrix},$$

then the convolution in AI is calculated as $ae + bf + cg + dh$ by the sweeping. On the other hand,

$$AB^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & g \\ f & h \end{pmatrix} = \begin{pmatrix} ae + bf & ag + bh \\ ce + df & cg + dh \end{pmatrix},$$

and

$$tr(AB^T) = ae + bf + cg + dh$$

for T is the transpose and tr is the trace. This is the same result as in AI.

3. Generalized Continuous Form of Matrix Expression of Convolution

If the matrix representing a function f is A and the matrix representing a function g is B, then the convolution of the functions f and g can be denoted by $tr(AB^T)$. Intuitively, the diagonal part of B^T corresponds to a graph of $g(t - \tau)$. The overlapping part of the graph can be interpreted as the concept of intersection, that is, the concept of multiplication. Thus, the generalized continuous form of the convolution in AI can be represented in a variant of Laplace-type transform given by

$$G_\alpha(f) = G(f) = u^\alpha \int_0^\infty e^{-\frac{t}{u}} f(t) dt$$

$$= u^\alpha \int_0^\infty \lim_{\delta \to 1} \left(1 + \frac{\delta - 1}{u}\right)^{-\frac{t}{\delta - 1}} f(t) dt.$$

If $f(t)$ is a function defined for all $t \geq 0$, an integral of Laplace-type transform $vi(f)$ is given by

$$F(\Delta) = vi(f) = \int_0^\infty e^{-t\Delta(\delta, u)} f(t) \, dt$$

$$= \int_0^\infty \left[1 + \frac{\delta - 1}{u}\right]^{-\frac{t}{\delta - 1}} f(t) \, dt$$

for $\delta > 1$ with

$$\Delta(\delta, u) = \frac{\ln\left[1 + \frac{\delta - 1}{u}\right]}{\delta - 1}.$$

Additionally, let $\Phi(u)$ be an arbitrary bounded function and let $V(f)$ be a variant of Laplace-type transform of $f(t)$. If $f(t)$ is a function defined for all $t \geq 0$, $V(f)$ is defined by

$$V(f) = \Phi(u) \int_0^\infty e^{-t\Delta} f(t) \, dt$$

$$= \Phi(u) \int_0^\infty \left[1 + \frac{\delta - 1}{u}\right]^{-\frac{t}{\delta - 1}} f(t) \, dt$$

for $\delta > 1$ with $\Delta = \Delta(\delta, u)$.

Based on the above two definitions, it is clear that the above variant of Laplace-type transform is represented as $V(f) = \Phi(u) \cdot vi(f)$ for an arbitrary function $\Phi(u)$. If so, let us see the relation with other integral transforms. Since

$$V(f) = \Phi(u) \int_0^\infty e^{-t\Delta} f(t) \, dt,$$

if $\delta \to 1+$ and $\Phi(u) = u^\alpha$, then it corresponds to the G_α transform. When we take $\delta \to 1$, $\Phi(u) = 1$, and $u = 1/s$, we get the Laplace transform. Similarly, when we take $\delta \to 1$, $\Phi(u) = u^{-1}$ ($\delta \to 1$, $\Phi(u) = u$), we get the Sumudu transform (Elzaki transform), respectively. In order to obtain a simple form of generalization, it is better to set $\phi(u)$ to u^α for an arbitrary integer α. However, it is judged that $\phi(u)$ is better than u^α as a suitable generalization, where $\phi(u)$ is a bounded arbitrary function. The reason is that $\phi(u)$ can express more integral transforms.

Lemma 1. *(Lebesgue dominated convergence theorem [16,17]). Let (X, M, μ) be a measure space and suppose $\{f_n\}$ is a sequence of extended real-valued measurable functions defined on X such that*
  (a) $\lim_{n\to\infty} f_n(x) = f(x)$ exists μ-a.e.
  (b) There is an integrable function g so that for each n, $|f_n| \le g$ μ-a.e.
  Then, f is integrable and

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Beppo Livi's theorem is a special form of Lemma 1. Its contents are as follows:

$$\int \sum_{n=1}^\infty g_n \, d\mu = \sum_{n=1}^\infty \int g_n \, d\mu$$

for (g_n) is a nondecreasing sequence. The details are can be found on page 71 in [16]. Note that the convolution of f and g is given by

$$(f * g)(t) = \int_0^\infty f(\tau) g(t - \tau) \, d\tau.$$

The following theorem is as follows. Since the proof is not difficult, we would like to cover just a few.

Theorem 1.

(1) *(Duality with Laplace transform) If $£(f) = F_*(s)$ is the Laplace transform of a function $f(t)$, then it satisfies the relation of $V(f) = \Phi(u) \cdot F_*(\Delta)$.*

(2) *(Shifting theorem) If $f(t)$ has the transform $F(u)$, then $e^{at} f(t)$ has the transform $\Phi(u) \cdot F(\Delta - a)$. That is,*

$$V[e^{at} f(t)] = \Phi(u) \cdot F(\Delta - a).$$

Moreover, If $f(t)$ has the transform $F(u)$, then the shifted function $f(t - a)h(t - a)$ has the transform $e^{-a\Delta} \cdot \Phi(u)F(\Delta)$. In formula,

$$V[f(t - a)h(t - a)] = e^{-a\Delta} \cdot \Phi(u)F(\Delta)$$

for $h(t - a)$ is Heaviside function (We write h since we need u to denote u-space).

(3) *(Linearity) Let $V(f)$ be the variant of Laplace-type transform. Then $V(f)$ is a linear operation.*

(4) *(Existence) If $f(t)$ is defined, piecewise continuous on every finite interval on the semi-axis $t \geq 0$ and satisfies*

$$|f(t)| \leq Me^{kt}$$

for all $t \geq 0$ and some constants M and k, then the variant of Laplace-type transform $V(f)$ exists for all $\Delta > k$.

(5) *(Uniqueness) If the variant of Laplace-type transform of a given function exists, then it is uniquely determined.*

(6) *(Heaviside function)*

$$vi[\,h(t-a)] = \int_0^\infty e^{-t\Delta} h(t-a)\, dt = \int_a^\infty e^{-t\Delta} \cdot 1\, dt$$

$$= e^{-a\Delta}/\Delta,$$

where h is Heaviside function.

(7) *(Dirac's delta function) We consider the function*

$$f_k(t-a) = \begin{cases} 1/k & \text{if } a \leq t \leq a+k \\ 0 & \text{otherwise.} \end{cases}$$

In a similar way to Heaviside, taking the integral of Laplace-type transform, we get

$$vi[f_k(t-a)] = \int_0^\infty e^{-t\Delta} f_k(t-a)\, dt = -\frac{1}{k\Delta}[e^{-t\Delta}]_a^{a+k}$$

$$= -\frac{1}{k\Delta}(e^{-(a+k)\Delta} - e^{-a\Delta}) = -\frac{1}{k\Delta}e^{-a\Delta}(e^{-k\Delta} - 1).$$

If we denote the limit of f_k as $\delta(t-a)$, then

$$vi(\delta(t-a)) = \lim_{k \to 0}\ (vi)[f_k(t-a)] = e^{-a\Delta}.$$

(8) *(Shifted data problems) For a given differential equation $y'' + ay' + by = r(t)$ subject to $y(t_0) = c_0$ and $y'(t_0) = c_1$, where $t_0 \neq 0$ and a and b are constant, we can set $t = t_1 + t_0$. Then $t = t_0$ gives $t_1 = 0$ and so, we have*

$$y_1'' + ay_1' + by_1 = r(t_1 + t_0),\ y_1(0) = c_0,\ y_1'(0) = c_1$$

for input $r(t)$. Taking the variant, we can obtain the output $y(t)$.

(9) *(Transforms of derivatives and integrals) Let a function f is n-th differentiable and integrable, and let us consider the fraction Δ as an operator. Then $V(f)$ of the n-th derivatives of $f(t)$ satisfies*

$$vi(f^{(n)}) = \Delta^n vi(f) - \sum_{k=1}^n \Delta^{n-k} f^{(k-1)}(0) \tag{2}$$

and

$$V\left[\int_0^t f(\tau)\, d\tau\right] = \Phi(u) \cdot \frac{1}{\Delta} V(f).$$

(10) *(Convolution) If two functions f and g are integrable for $*$ is the convolution, then $V(f * g)$ satisfies*

$$V(f * g) = \Phi(u) \cdot F(\Delta)G(\Delta)$$

for $V(f) = F(\Delta)$.

Proof. (5) Assume that $V(f)$ exists by $V(f_1)$ and $V(f_2)$ both. If $V(f_1) \neq V(f_2)$ for $f_1 = f_2$, then

$$
\begin{aligned}
V(f_1) - V(f_2) &= \Phi(u) \int_0^\infty e^{-t\Delta} f_1(t)\, dt - \Phi(u) \int_0^\infty e^{-t\Delta} f_2(t)\, dt \\
&= \Phi(u) \int_0^\infty e^{-t\Delta}(f_1(t) - f_2(t))\, dt \\
&= V(f_1 - f_2) = 0.
\end{aligned}
$$

This is a contradiction on $V(f_1) \neq V(f_2)$, and hence the transform is uniquely determined. Conversely, if two functions f_1 and f_2 have the same transform (i.e., if $V(f_1) = V(f_2)$), then

$$
V(f_1) - V(f_2) = \Phi(u) \int_0^\infty e^{-t\Delta}(f_1(t) - f_2(t))\, dt = 0,
$$

and so $f_1 = f_2$ a.e. Hence $f_1 = f_2$ excepting for the set of measure zero.

(9) Note that $vi(f) = \int_0^\infty e^{-t\Delta} f(t)\, dt$, and let us approach the proof by induction. In case of $n = 1$,

$$
vi(f') = \int_0^\infty e^{-t\Delta} f'(t)\, dt.
$$

Integrating by parts, we have

$$
vi(f') = [e^{-t\Delta} f(t)]_0^\infty + \Delta \int_0^\infty e^{-t\Delta} f(t)\, dt
$$

$$
= -f(0) + \Delta vi(f).
$$

which is true by (2).

Next, let us suppose that $n = m$ is valid for some m. Thus,

$$
vi(f^{(m)}) = \Delta^m vi(f) - \sum_{k=1}^m \Delta^{m-k} f^{(k-1)}(0)
$$

holds for $f^{(m)}$ is the m-th derivative of f. Let us show that

$$
vi(f^{(m+1)}) = \Delta^{m+1} vi(f) - \sum_{k=1}^{m+1} \Delta^{m+1-k} f^{(k-1)}(0). \tag{3}
$$

Now we start with the left-hand side of (2).

$$
vi(f^{(m+1)}) = \int_0^\infty e^{-t\Delta} f^{(m)}(t)\, dt
$$

$$
= \Delta vi(f^{(m)}) - f^{(m)}(0)
$$

$$
= \Delta[\Delta^m vi(f) - \sum_{k=1}^m \Delta^{m-k} f^{(k-1)}(0)] - f^{(m)}(0)
$$

$$
= \Delta^{m+1} vi(f) - \sum_{k=1}^m \Delta^{m+1-k} f^{(k-1)}(0)] - f^{(m)}(0)
$$

$$
= \Delta^{m+1} vi(f) - \sum_{k=1}^{m+1} \Delta^{m+1-k} f^{(k-1)}(0).
$$

Therefore, this theorem is valid for an arbitrary natural number n. Putting $g(t) = \int_0^t f(\tau)d\tau$,

$$
vi(f(t)) = vi(g'(t)) = \Delta vi(g) - g(0) = \Delta vi(g)
$$

follows.　□

As the direct results of (9), $vi(f') = \Delta vi(f) - f(0)$ and $vi(f'') = \Delta^2 vi(f) - \Delta f(0) - f'(0)$ are follow.

For example, we consider $y'' - y = t$ subject to $y(0) = 1$ and $y'(0) = 1$. Taking the integral of Laplace-type transform on both sides, we have

$$\Delta^2 Y - \Delta y(0) - y'(0) - Y = 1/\Delta^2$$

for $Y = (vi)(y)$. Organizing this equation, we get $(\Delta^2 - 1)Y = \Delta + 1 + 1/\Delta^2$. Simplification gives

$$Y = \frac{1}{\Delta - 1} + \frac{1}{\Delta^2 - 1} - \frac{1}{\Delta^2}.$$

From the relation of $V(f) = \Phi(u) \cdot F_*(\Delta)$, we have the solution

$$y(t) = -t + 2 \sin ht + \cos ht = e^t + \sin ht - t,$$

where h is hyperbolic function.

Example 3. *(Integral equations of Volterra type) Find the solution of*

$$(1)\ y(t) + \int_0^t (t - \tau) y(\tau) d\tau = 1.$$

$$(2)\ y(t) - \int_0^t y(\tau) \sin(t - \tau) d\tau = t.$$

$$(3)\ y(t) - \int_0^t (1 + \tau) y(t - \tau) d\tau = 1 - \sinh t.$$

Solution.

(1) Since this equation is $y + y * t = 1$, taking the integral of Laplace-type transform on both sides, we have

$$Y + (Y \cdot \frac{1}{\Delta^2}) = \frac{1}{\Delta}$$

for $Y = vi(y)$. Thus

$$Y = \frac{\Delta}{\Delta^2 + 1}$$

and so, we obtain the solution $y = \cos t$.

Let us do the check by expansion. Expanding, we get $y''(t) + y(t) = 0$. Since $\int_a^a f = 0$, we get $y(0) = 1$ and $y'(0) = 0$. Thus, we obtain $y = \cos t$.

(2) This is rewritten as a convolution

$$y(t) - y * \sin t = t.$$

Taking the integral of Laplace-type transform, we have

$$Y(u) - Y(u)\frac{1}{\Delta^2 + 1} = Y(u)(1 - \frac{1}{\Delta^2 + 1}) = \frac{1}{\Delta^2}$$

for $Y = vi(y)$. The solution is

$$Y(u) = \frac{1}{\Delta^2} + \frac{1}{\Delta^4}$$

and gives the answer

$$y(t) = t + \frac{1}{6}t^3.$$

(3) Note that the equation is the same as $y - (1+t) * y = 1 - \sinh t$. Taking the transform, we get

$$Y - \left(\frac{1}{\Delta} + \frac{1}{\Delta^2}\right)Y = \frac{1}{\Delta} - \frac{1}{\Delta^2 - 1},$$

and hence

$$Y\left(1 - \frac{1}{\Delta} - \frac{1}{\Delta^2}\right) = \frac{\Delta^2 - \Delta - 1}{\Delta(\Delta^2 - 1)}.$$

Simplification gives

$$Y = \frac{\Delta}{\Delta^2 - 1},$$

and so, we obtain the answer

$$y(t) = \cosh t$$

by the relation of $V(f) = \Phi(u) \cdot F_*(\Delta)$.

Let us turn the topic to initial value problem of the convolution. The initial value problem

$$ay'' + by' + cy = f(t), \quad y(0) = y_0, \quad y'(0) = y_0'$$

gives

$$(a\Delta^2 + b\Delta + c)Y(\Delta) - (a\Delta + b)y(0) - ay'(0) = F(\Delta),$$

where $Y(\Delta) = vi(y)$ and $F(\Delta) = vi(f)$. Simplification gives

$$Y(\Delta) = \frac{1}{a\Delta^2 + b\Delta + c} \cdot F(\Delta) + y_0 \cdot \frac{a\Delta + b}{a\Delta^2 + b\Delta + c}$$

$$+ y_0' \cdot \frac{a}{a\Delta^2 + b\Delta + c}.$$

If we put the system function $H(\Delta) = (a\Delta^2 + b\Delta + c)^{-1}$, then

$$Y(\Delta) = H(\Delta)F(\Delta) + y_0(a\Delta + b)H(\Delta) + y_0' a H(\Delta).$$

Since $H(\Delta)F(\Delta) = vi(h)vi(f) = vi(h * f)$ for $H(\Delta) = vi(h(t))$, taking the inverse transform, we have

$$y = (h * f) + y_0 \, vi^{-1}\{(a\Delta + b)H(\Delta)\}$$

$$+ y_0' \, vi^{-1}\{aH(\Delta)\}.$$

Theorem 2. *(Differentiation and integration of transforms) Let us put $V = F(\Delta)$ and $Y = V(y)$. Then*

(1) $\quad V(t^n f(t)) = \Phi(u) \cdot (-1)^n \dfrac{d^n}{ds^n} F(\Delta)$

(2) $\quad V(f(t)/t) = \Phi(u) \displaystyle\int_{\Delta}^{\infty} F(\delta)\,d\delta.$

Proof. This is an immediate consequence of $V(f) = \Phi(u) \cdot F_*(\Delta)$ and $V(f) = \Phi(u) \cdot vi(f)$. For this reason, detailed proofs are omitted.

The statements below are the immediate results of Theorem 2.

(1) $\quad V(tf(t)) = -\Phi(u)F'(\Delta)$

(2) $\quad V(tf'(t)) = -\Phi(u)\left(Y + \Delta \dfrac{dY}{d\Delta}\right)$

(3) $\quad V(tf''(t)) = \Phi(u) \cdot \left(-2\Delta Y - \Delta^2 \dfrac{dY}{d\Delta} + y(0)\right)$

Let us check examples for temperature in an infinite bar and displacement in a semi-infinite string by the variant of Laplace-type transform. □

Example 4. *(Semi-infinite string) Find the displacement $w(x,t)$ of an elastic string subject to the following conditions [3].*

(a) *The string is initially at rest on the x-axis from $x = 0$ to ∞.*
(b) *For $t > 0$ the left end of the string is moved in a given fashion, namely, according to a single sine wave*

$$w(0,t) = f(t) = \begin{cases} \sin 2t & \text{if } 0 \le t \le \pi \\ 0 & \text{otherwise.} \end{cases}$$

(c) *Furthermore, $w(x,t) \to 0$ as $x \to \infty$ for $t \ge 0$.*

Then the displacement w is

$$w(x,t) = f(t - \frac{x}{c})h(t - \frac{x}{c}) = \begin{cases} \sin 2(t - \frac{x}{c}) & \text{if } x/c < t < x/c + \pi \\ 0 & \text{otherwise,} \end{cases}$$

where h is Heaviside function.

The proof is simple, and the interchangeability of limit and integral in the proof process guarantees its validity by the Lebesgue dominated convergence theorem.

Example 5. *(Temperature in an infinite bar) Find the temperature w in an infinite bar if the initial temperature is*

$$f(x) = w(x,0) = \begin{cases} k_0 \ (\text{constant}) & \text{if } |x| < 1 \\ 0 & \text{otherwise} \end{cases}$$

with $w(0,t) = 0$.

Solution. Taking the integral of Laplace-type transform on both sides of $w_t = c^2 w_{xx}$, we have

$$\Delta F - w(x,0) = c^2 \frac{\partial^2 F}{\partial x^2}$$

for $F(x,u) = vi[w(x,t)]$. Organizing the equality, we get

$$\frac{\partial^2 F}{\partial x^2} - \frac{\Delta}{c^2}F = -\frac{1}{c^2}w(x,0).$$

Organizing this equality, we get

$$F(x,u) = A(u)e^{-\sqrt{\Delta}x/c} + B(u)e^{\sqrt{\Delta}x/c}$$

$$+ e^{-\sqrt{\Delta}x/c} \int \frac{c}{2\sqrt{\Delta}} e^{\sqrt{\Delta}x/c} \frac{1}{c^2} w(x,0) \, dx$$

$$- e^{\sqrt{\Delta}x/c} \int \frac{c}{2\sqrt{\Delta}} e^{-\sqrt{\Delta}x/c} \frac{1}{c^2} w(x,0) \, dx, \tag{4}$$

where the Wronskian $W = 2\sqrt{\Delta}/c$. The value $\lim_{x\to\infty} f(x) = 0$ gives $\lim_{x\to\infty} F(x,u) = 0$, and hence $B(u) = 0$. Thus, from (4), we get

$$F(x,u) = A(u)e^{-\sqrt{\Delta}x/c}$$

$$+\frac{w(x,0)}{2c\sqrt{\Delta}}\left(e^{-\sqrt{\Delta}x/c}\int e^{\sqrt{\Delta}x/c}\,dx - e^{\sqrt{\Delta}x/c}\int e^{-\sqrt{\Delta}x/c}\,dx\right).$$

By the direct calculation, we have

$$F(x,u) = A(u)e^{-\sqrt{\Delta}x/c} + \frac{w(x,0)}{\Delta}.$$

From the formula of $vi(f) = F(\Delta)$ for $F(s) = \pounds(f)$ and $s = \Delta$, we know

$$vi\left(\frac{k}{2\sqrt{\pi t^3}}\,e^{-\frac{k^2}{4t}}\right) = e^{-k\sqrt{\Delta}}\ (k > 0)$$

and $vi(1) = 1/\Delta$. Taking the inverse transform, we obtain the temperature $w(x,t)$ as follows:

$$w(x,t) = A(t) * \frac{x}{2c\sqrt{\pi t^3}}\,e^{-\frac{x^2}{4c^2 t}} + k_0$$

on $|x| < 1$, and $*$ is the convolution. In case of $|x| > 1$, we have the solution

$$w(x,t) = A(t) * \frac{x}{2c\sqrt{\pi t^3}}\,e^{-\frac{x^2}{4c^2 t}}.$$

In the above equality, we note that

$$vi^{-1}\left[A(u)e^{-\sqrt{\Delta}x/c}\right] = vi^{-1}\left[vi(A(t)) \cdot vi\left(\frac{x}{2c\sqrt{\pi t^3}}\,e^{-\frac{x^2}{4c^2 t}}\right)\right]$$

$$= vi^{-1}\left[vi\left\{A(t) * \frac{x}{2c\sqrt{\pi t^3}}\,e^{-\frac{x^2}{4c^2 t}}\right\}\right] = A(t) * \frac{x}{2c\sqrt{\pi t^3}}\,e^{-\frac{x^2}{4c^2 t}}$$

because $vi(f * g) = F(\Delta)g(\Delta)$ for $vi(f) = F(\Delta)$.

4. Conclusions

In this study, the concept of convolution in convolutional neural networks (CNN) was presented mathematically and tried to connect with the concept of convolution in mathematics. As a continuous form of convolution in CNN, a new form of Laplace-type transform has been proposed. In the future, we will study the change of convolution in CNN by changing the stride. In addition to this, we shall also explore the possibility of our applying our newly defined Laplace-type transform in obtaining certain new and interesting results involving generalized hypergeometric functions that would certainly unify and generalized the results available in the literature and may be potentially useful from an applications point of view.

Author Contributions: Conceptualization, A.K.R.; validation, Y.H.G.; and writing, H.K. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*. [CrossRef] [PubMed]
2. Kim, H. The intrinsic structure and properties of Laplace-typed integral transforms. *Math. Probl. Eng.* **2017**, *2017*, 1–8. [CrossRef]

3.   Kreyszig, E. *Advanced Engineering Mathematics*; Wiley: Singapore, 2013.

4.   Watugula, G.K. Sumudu Transform: A new integral transform to solve differential equations and control engineering problems. *Integr. Educ.* **1993**, *24*, 35–43. [CrossRef]

5.   Elzaki, T.M.; Ezaki, S.M.; Hilal, E.M.A. ELzaki and Sumudu Transform for Solving some Differential Equations. *Glob. J. Pure Appl. Math.* **2012**, *8*, 167–173.

6.   Mohand, M.; Mahgoub, A. The New Integral Transform 'Mohand Transform. *Adv. Theor. Appl. Math.* **2017**, *12*, 113–120.

7.   Belgacem, F.B.M.; Silambarasan, R. Theory of natural transform. *Math. Eng. Sci. Aerosp.* **2012**, *3*, 105–135.

8.   Bertrand, J.; Bertrand, P.; Ovarlez, J.P. *The Mellin Transform, The Transforms and Applications*; Poularkas, A.D., Ed.; CRC Press: Boca Raton, FL, USA, 1996.

9.   Jhanthanam, S.; Nonlaopon, K.; Orankitjaroen, S. Generalized Solutions of the Third-Order Cauchy-Euler Equation in the Space of Right-Sided Distributions via Laplace Transform. *Mathematics* **2019**, *7*, 376. [CrossRef]

10. Kim, H. The solution of the heat equation without boundary conditions. *Dyn. Syst. Appl.* **2018**, *27*, 653–662.

11. Supaknaree, S.; Nonlaopon, K.; Kim, H. Further properties of Laplace-type integral transforms. *Dyn. Syst. Appl.* **2019**, *28*, 195–215.

12. Koepf, W.; Kim, I.; Rathie, A.K. On a New Class of Laplace-Type Integrals Involving Generalized Hypergeometric Functions. *Axioms* **2019**, *8*, 87. [CrossRef]

13. Sung, T.; Kim, I.; Rathie, A.K. On a new class of Eulerian's type integrals involving generalized hypergeometric functions. *Aust. J. Math. Anal. Appl.* **2019**, *16*, 1–15.

14. Noeiaghdam, S.; Fariborzi Araghi, M.A.; Abbasbandy, S. Finding optimal convergence control parameter in the homotopy analysis method to solve integral equations based on the stochastic arithmetic. *Numer. Algorithms* **2019**, *81*. [CrossRef]

15. Negnevitsky, M. *Artificial Intelligence*; Addison-Wesley: Essex, England, 2005.

16. Cohn, D.L. *Measure Theory*; Birkhäuser: Boston, MA, USA, 1980.

17. Jang, J.; Kim, H. An application of monotone convergence theorem in PDEs and Fourier analysis. *Far East J. Math. Sci.* **2015**, *98*, 665–669.