

基于SVM的手写数字识别

马 赫*, 崔橙阳, 贾文硕, 张政华, 李文臣

西京学院计算机学院, 陕西 西安

收稿日期: 2025年3月13日; 录用日期: 2025年4月30日; 发布日期: 2025年5月12日

摘 要

本文主要研究基于支持向量机(SVM)的手写数字识别方法。通过对SVM算法的原理进行深入分析, 结合手写数字图像的特征提取方法, 构建了一个高效的手写数字识别系统。初始时, 固定惩罚参数 C (选用0.1) 和 γ 值(选用scale), 采用RBF核和linear核函数进行训练, 最后对RBF核函数和linear核函数的各项性能进行比较, 如召回率、准确率等, 选取针对手写数字识别这一场景下最优的核函数。最后实验结果表明, RBF核函数的各项性能指标如准确率、召回率、F1函数等值在这种场景下要优于linear核函数。最后选取RBF核函数, 再进一步调整惩罚参数 C 和 γ 值以训练出最优的模型。

关键词

支持向量机(SVM), 手写数字识别, RBF核函数, Linear核函数

Handwritten Digit Recognition Based on SVM

He Ma*, Chengyang Cui, Wenshuo Jia, Zhenghua Zhang, Wenchen Li

School of Computer Science, Xijing University, Xi'an Shaanxi

Received: Mar. 13th, 2025; accepted: Apr. 30th, 2025; published: May 12th, 2025

Abstract

This paper mainly studies the handwritten numeral recognition method based on Support Vector Machine (SVM). Through the in-depth analysis of the principle of SVM algorithm, combined with the feature extraction method of handwritten digital images, an efficient handwritten digital recognition system is constructed. At the beginning, the penalty parameters C (0.1) and γ (scale) are fixed, and RBF kernel and linear kernel function are used for training. Finally, the performances of RBF kernel function and linear kernel function, such as recall rate and accuracy rate, are compared, and the optimal kernel function for handwritten numeral recognition is selected. Finally, the experimental results show that the

*通讯作者。

文章引用: 马赫, 崔橙阳, 贾文硕, 张政华, 李文臣. 基于 SVM 的手写数字识别[J]. 人工智能与机器人研究, 2025, 14(3): 501-509. DOI: 10.12677/airr.2025.143049

performance indexes of RBF kernel function, such as accuracy, recall and F1 function equivalence, are better than linear kernel function in this scenario. Finally, RBF kernel function is selected, and then the penalty parameters C and γ are further adjusted to train the optimal model.

Keywords

Support Vector Machines (SVM), Handwritten Digit Recognition, RBF Kernel Function, Linear Kernel Function

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着信息技术的飞速发展, 手写数字识别作为模式识别领域的一个重要研究课题, 已在多个实际应用中得到广泛应用, 例如自动邮政编码识别、银行支票数字识别以及表格数据的自动录入等。然而, 手写数字识别面临着诸如样本变形、噪声干扰以及多样化的书写风格等挑战, 这使得传统的模式识别方法在识别准确率和泛化能力上存在一定的局限性[1]。为了提高手写数字识别的效果, 研究者们不断探索更加高效和准确的算法。

支持向量机(Support Vector Machine, SVM)作为一种强大的监督学习算法, 已广泛应用于各种模式识别任务中。SVM 能够有效处理高维数据, 特别是在处理小样本学习问题时, 展现了其独特的优势。通过选择合适的核函数和惩罚参数, SVM 能够构建出精确的分类模型, 尤其适用于手写数字识别这样的复杂模式识别任务[2]。为了进一步提高识别性能, 许多研究将 SVM 与特征提取方法结合, 构建了多种手写数字识别系统。

本文的研究聚焦于基于 SVM 的手写数字识别方法。我们通过深入分析 SVM 的工作原理, 并结合手写数字图像的特征提取方法[3], 设计了一个高效的手写数字识别系统。首先, 选择固定的惩罚参数 C (设为 0.1) 和 γ 值 (设为 scale), 并使用 RBF 核和线性核函数对模型进行训练。在训练过程中, 我们比较了 RBF 核函数和线性核函数在手写数字识别任务中的性能, 包括准确率、召回率、F1 值等关键性能指标。实验结果表明, RBF 核函数在各项性能指标上优于线性核函数。因此, 本文选取了 RBF 核函数, 并进一步优化惩罚参数 C 和 γ 值, 训练出最优的手写数字识别模型。

在国内, 手写数字识别的研究已经取得了显著进展, 研究者们不仅关注传统的机器学习方法(如 SVM、决策树等), 还逐步探索深度学习技术的应用[4][5]。基于卷积神经网络(CNN)的方法在这一领域得到了广泛关注, 并取得了较好的效果。然而, SVM 在小样本情况下仍然展现出较高的识别精度, 并具有较强的泛化能力, 因此, 在手写数字识别的应用中仍占据重要地位。国内许多高校和科研机构在这一领域开展了大量的研究, 推动了相关技术的本土化发展, 并取得了诸多创新成果[6]。

在国外, 手写数字识别的研究历史悠久, 许多经典算法和模型相继出现。例如, LeNet-5 卷积神经网络的提出标志着深度学习在手写数字识别中的应用取得了突破[7]。近年来, 深度学习技术的发展使得基于神经网络的方法在准确率上不断刷新纪录, 但 SVM 作为一种经典的机器学习方法, 依然在特定任务和数据条件下表现出强大的竞争力[8]。

本文的贡献在于系统地探讨了基于 SVM 的手写数字识别方法, 并对比了不同核函数在该任务中的表现, 提出了最适合手写数字识别任务的 RBF 核函数。通过对参数进行调优, 进一步提高了识别准确性, 为今后的研究提供了可行的参考。

2. 算法原理

支持向量机(SVM)是一种强大的监督学习算法,用于分类和回归任务。它的核心思想是找到一个最优超平面,以最大化不同类别之间的间隔[9]。在一维 SVM 通过找到一个点来划分不同的样本;在二维 SVM 通过找到一条线来对不同的样本进行划分。而对于一些简单的样本, SVM 可以很迅速地找到一条直线来划分不同的样本,这种情况被称为线性可分的。而对于一些复杂的样本点则需要通过一条曲线来划分样本点,这种情况被称为线性不可分的[10]。在二维空间中,这条用于划分不同样本点的线就被称为超平面;距离这条线最近的点就被称为支持向量。

2.1. 线性可分情况下的 SVM

在二维空间中,如果存在一个线性函数能够将两类不同的样本完全分开,那么这个问题就是线性可分的[11]。SVM 的基本思想是寻找一个最优的超平面,使得两类样本之间的间隔最大。这个超平面可以表示为:

$$w^T x + b = 0$$

其中, w^T 是超平面的法向量, b 是截距, x 是输入样本。

对于线性可分的情况,通过求解以下优化问题可以得到最优超平面:

$$\begin{aligned} \min_{w,b} \quad & \|w\|^2 \\ \text{subject to} \quad & y_i (w^T x_i + b) \geq 1 \end{aligned}$$

其中, $i = 1, 2, \dots, n$, n 是样本数量, y_i 是样本的类别标签,取值为 ± 1 。

根据拉格朗日对偶性,将上述优化问题转化为对偶问题进行求解:

$$\begin{aligned} \max_a \quad & \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_i a_j y_i y_j (x_i^T x_j) \\ \text{subject to} \quad & \sum_{i=1}^n a_i y_i, a_i \geq 0 \end{aligned}$$

其中, $i = 1, 2, \dots, n$ 。最后求解得到最优解:

$$a^* = (a_1^*, a_2^*, \dots, a_n^*)$$

则最优超平面的法向量

$$w^* = \sum_{i=1}^n a_i^* y_i x_i$$

则最优超平面的截距

$$b^* = y_i - \sum_{i=1}^n a_i^* y_i (x_i^T x_j)$$

其中, $a_j^* > 0$ 对应的样本点 (x_j, y_j) 称为支持向量。

2.2. 线性不可分情况下的 SVM

在实际应用中,很多问题往往是线性不可分的。为了解决这个问题, SVM 引入了核函数的概念。核函数将低维空间中的线性不可分问题映射到高维空间中,使其在高维空间中变得线性可分[12] [13]。

设原始空间为 X , 映射后的高维空间为 H , 核函数

$$K(x_i, y_j) = \phi(x_i)^T \phi(x_j)$$

其中, $\phi(x)$ 是将 x 从原始空间映射到高维空间的函数。

则在高维空间中的优化问题为:

$$\begin{aligned} \min_a \quad & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_i a_j y_i y_j K(x_i, x_j) \\ \text{subject to} \quad & \sum_{i=1}^n a_i y_i = 0, 0 \leq a_i \leq C, i = 1, 2, \dots, n \end{aligned}$$

其中, C 是惩罚参数。

3. 算法实现

3.1. 数据集选择

MNIST 数据集来自美国国家标准与技术研究所(National Institute of Standards and Technology, NIST) [3]。训练集(training set)由来自 250 个不同人手写的数字构成, 其中 50%是高中学生, 50%来自人口普查局(the Census Bureau)的工作人员。测试集(test set)也是同样比例的手写数字数据, 但保证了测试集和训练集的作者集不相交。

因此, 本实验使用 MNIST 手写数字数据集进行实验, 该数据集包含 60,000 个训练样本和 10,000 个测试样本, 每个样本都是一个 28×28 的灰度图像[14]。

3.2. 不同核函数的性能比较

本文分别使用了高斯核函数(RBF)和线性核函数(linear)来对 SVM 分类器进行训练和测试, 比较不同核函数的准确率、召回率等性能指标。实验结果表明, 高斯核函数在手写数字识别中表现出了最好的性能。

设线性核函数为:

$$K(x, y) = x \cdot y + c$$

高斯核函数为:

$$K(x, y) = e^{-\gamma \|x - y\|^2}$$

其中, c, d, γ 为参数。

1) RBF 核函数

在 RBF 核函数训练中, 惩罚参数 C 选取为 0.1, γ 值选取为 scale, 代码如下:

```
param_grid = {
    'C': [0.1],
    'gamma': ['scale'],
    'kernel': ['rbf'],
}
```

通过五折交叉法训练, 每一折都尝试作为验证集, 最后计算出的评分分别为 0.957 (第一折)、0.954 (第二折)、0.951 (第三折)、0.952 (第四折)、0.959 (第五折), 训练结果如图 1 所示。

```
Fitting 5 folds for each of 1 candidates, totalling 5 fits
[CV 4/5] END ....C=0.1, gamma=scale, kernel=rbf;; score=0.952 total time=15.5min
[CV 3/5] END ....C=0.1, gamma=scale, kernel=rbf;; score=0.951 total time=15.5min
[CV 2/5] END ....C=0.1, gamma=scale, kernel=rbf;; score=0.954 total time=15.5min
[CV 1/5] END ....C=0.1, gamma=scale, kernel=rbf;; score=0.957 total time=15.6min
[CV 5/5] END ....C=0.1, gamma=scale, kernel=rbf;; score=0.959 total time=15.6min
```

Figure 1. RBF kernel function five-fold crossover model performance evaluation
图 1. RBF 核函数五折交叉模型成绩评估

训练准确率、召回率、f1 分数以及满足的图片数量如图 2 所示。

```
Best cross-validation score: 0.9545666666666668
Classification Report:
```

	precision	recall	f1-score	support
0	0.97	0.99	0.98	980
1	0.98	0.99	0.98	1135
2	0.96	0.95	0.95	1032
3	0.95	0.96	0.96	1010
4	0.95	0.96	0.96	982
5	0.96	0.95	0.95	892
6	0.96	0.97	0.97	958
7	0.97	0.94	0.95	1028
8	0.95	0.95	0.95	974
9	0.95	0.93	0.94	1009
accuracy			0.96	10000
macro avg	0.96	0.96	0.96	10000
weighted avg	0.96	0.96	0.96	10000

Figure 2. Various performance indicators of RBF kernel function
图 2. RBF 核函数各种性能指标

混淆矩阵(Confusion Matrix)同样也是一种用于评估分类模型性能的工具，特别是在多分类问题中。它显示了模型预测结果与真实标签之间的关系，该模型的混淆矩阵如图 3 所示。

```
Confusion Matrix:
```

[	969	0	2	0	0	4	3	1	1	0]
[	0	1122	3	2	0	1	3	1	3	0]
[	6	0	982	8	8	0	5	9	14	0]
[	0	0	9	967	0	12	0	9	10	3]
[	1	1	3	0	947	0	7	2	2	19]
[	6	1	2	15	3	844	11	1	7	2]
[	9	3	2	0	4	6	931	0	3	0]
[	0	10	19	2	3	0	0	967	4	23]
[	4	0	3	10	7	11	6	6	924	3]
[	7	7	1	10	25	5	1	6	5	942]]

```
Model saved to svm_mnist_model.joblib
```

Figure 3. RBF kernel function confusion matrix
图 3. RBF 核函数混淆矩阵

最后将准备好的测试集进行测试，识别准确率达到 100%。结果如图 4 所示。

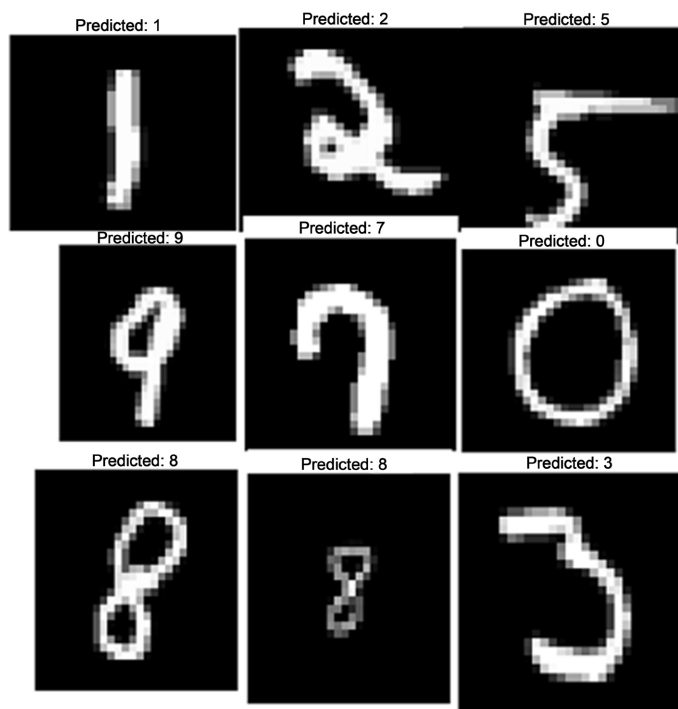


Figure 4. RBF kernel function test results
图 4. RBF 核函数测试结果

2) linear 核函数

在 linear 核函数训练中，同样地选取惩罚参数 C 为 0.1， γ 值为 scale，代码如下：

```
param_grid = {
    'C': [0.1],
    'gamma': ['scale'],
    'kernel': ['linear'],
}
```

通过五折交叉法训练[15]，每一折都作为验证集最后计算出的评分分别为 0.946 (第一折)、0.939 (第二折)、0.939 (第三折)、0.936 (第四折)、0.945 (第五折)，linear 核函数训练结果如图 5 所示。

```
Fitting 5 folds for each of 1 candidates, totalling 5 fits
[CV 4/5] END .C=0.1, gamma=scale, kernel=linear; score=0.936 total time= 6.2min
[CV 2/5] END .C=0.1, gamma=scale, kernel=linear; score=0.939 total time= 6.3min
[CV 3/5] END .C=0.1, gamma=scale, kernel=linear; score=0.939 total time= 6.3min
[CV 1/5] END .C=0.1, gamma=scale, kernel=linear; score=0.946 total time= 6.3min
[CV 5/5] END .C=0.1, gamma=scale, kernel=linear; score=0.945 total time= 6.3min
Grid search completed.
```

Figure 5. Linear kernel function training results
图 5. Linear 核函数训练结果

Linear 核函数对应的训练准确率、召回率、f1 分数以及满足的图片数量如图 6 所示。

Linear 核函数混淆矩阵如图 7 所示。

使用 Linear 核函数训练的模型测试时识别错误的一张图片，如图 8 所示。

Classification Report:

	precision	recall	f1-score	support
0	0.96	0.99	0.97	980
1	0.98	0.99	0.98	1135
2	0.94	0.94	0.94	1032
3	0.91	0.94	0.93	1010
4	0.94	0.96	0.95	982
5	0.92	0.90	0.91	892
6	0.96	0.96	0.96	958
7	0.95	0.94	0.95	1028
8	0.94	0.91	0.92	974
9	0.96	0.92	0.94	1009
accuracy			0.95	10000
macro avg	0.95	0.95	0.95	10000
weighted avg	0.95	0.95	0.95	10000

Figure 6. Various indicators of linear kernel function
图 6. Linear 核函数各种指标

Confusion Matrix:

```
[ [ 967  0  2  1  1  6  1  1  1  0]
 [  0 1124  2  2  0  1  2  1  3  0]
 [  9  2  974 10  5  2  8  8 14  0]
 [  2  0  14  953  0 16  0 11 13  1]
 [  1  1  6  0  942  0  6  2  2 22]
 [  8  3  1  34  5 807 12  1 19  2]
 [  6  2  8  1  4 15  921  1  0  0]
 [  2  5 23  8  8  0  0  968  0 14]
 [  4  5  5 24  6 27  6  6 888  3]
 [  6  6  2 11 31  2  1 15  7 928]]
```

Model saved to linear.joblib

Figure 7. Linear kernel function confusion matrix
图 7. Linear 核函数混淆矩阵

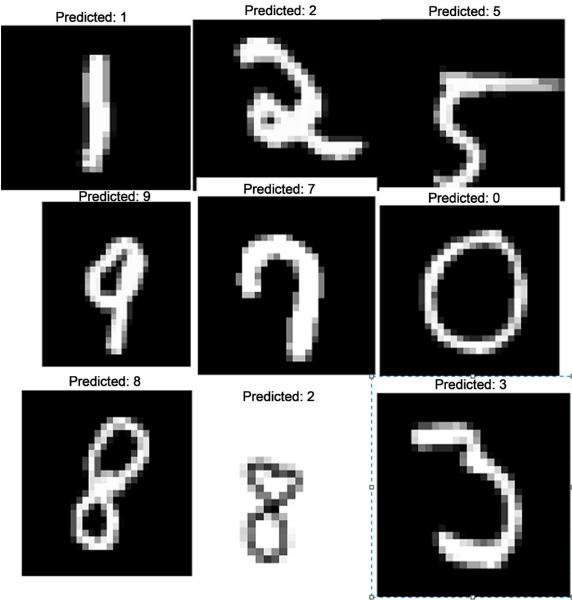


Figure 8. Linear kernel function test results
图 8. Linear 核函数测试结果

4. 总结

本文研究了基于支持向量机(SVM)的手写数字识别方法,通过深入分析 SVM 算法的原理,并结合手写数字图像的特征提取方法,构建了一个高效的手写数字识别系统。实验结果表明,该系统在多个性能指标上表现出较高的准确性和稳定性,尤其在识别准确率和召回率上取得了较为优异的成绩,证明了 SVM 在处理手写数字识别任务中的有效性和优势。通过比较不同核函数(RBF 核与线性核函数)的性能,发现 RBF 核在本研究的手写数字识别任务中优于线性核函数,尤其是在处理复杂、变形的手写体时,RBF 核能够更好地捕捉数据的非线性特征,提升了分类精度。

在系统的优化过程中,进一步调整了惩罚参数 C 和 γ 值,确保模型的高效性和稳定性。这一过程表明, SVM 不仅能够较小的样本集上达到较高的准确率,还能够通过合理的参数调整来优化模型表现。这些结果为基于 SVM 的手写数字识别提供了有力的实践依据,证明了 SVM 在模式识别任务中的强大适用性,尤其是在资源有限且要求高准确率的应用场景中[16]。

然而,尽管本研究在手写数字识别领域取得了令人满意的成果,但仍存在进一步优化的空间。未来的研究可以从以下几个方面进一步提高 SVM 分类器的性能。首先,在特征提取方面,当前的方法主要依赖手工特征,尽管取得了不错的效果,但可以探索更加自动化和智能化的特征学习方法,如结合深度学习技术,利用卷积神经网络(CNN)自动提取图像特征,从而提升特征表示的能力和识别精度。其次,尽管 SVM 在小样本学习和高维数据下表现良好,但在处理大规模数据集时,训练速度可能成为瓶颈。未来的研究可以考虑采用并行化处理和优化算法,进一步提高识别速度和计算效率。

此外,本研究的成果可以推广应用到更多的实际场景中,如自动化数据录入、智能表单识别、数字化档案处理等,进一步推动手写数字识别技术在各行各业的应用[17]。在商业、金融、教育等领域,手写数字识别技术的应用潜力巨大,可以有效提高效率、降低人工成本,并改善用户体验。

总的来说,基于 SVM 的手写数字识别方法在准确性、稳定性和应用潜力方面展现了良好的前景。随着技术的不断发展,结合其他先进的机器学习方法,特别是深度学习技术,将有可能进一步提升识别精度和速度,为实际应用提供更加高效和可靠的解决方案。

参考文献

- [1] 田春婷. 基于 PCA 降维的 MNIST 手写数字识别优化[J]. 现代信息科技, 2024, 8(16): 64-68.
- [2] 徐平, 徐海东, 杨拓, 等. 三层衍射神经网络实现手写数字识别[J]. 物理学报, 2022, 71(18): 137-144.
- [3] 叶佩, 江涛. 基于 BP 特征筛选的模糊聚类手写体数字识别[J]. 武汉理工大学学报, 2007, 29(6): 128-130.
- [4] 苗夺谦, 张红云, 李道国, 等. 基于主曲线的脱机手写数字识别[J]. 电子学报, 2005, 33(9): 1639-1643.
- [5] 林列, 常胜江. 用于手写体数字不变性特征提取及识别的自组织算法[J]. 电子学报, 2003, 31(10): 1506-1509.
- [6] 房孝猛, 王华来, 徐晖, 等. 基于 SVM 与近红外 TDLAS 技术的多组分痕量气体识别与检测[J]. 光谱学与光谱分析, 2024, 44(10): 2909-2915.
- [7] 张微微, 璩怡, 王强, 等. 同步荧光光谱技术结合支持向量机对掺杂牛奶智能判别研究[J]. 光谱学与光谱分析, 2024, 44(9): 2428-2433.
- [8] 龙泽凯, 陈聪飞, 郝东来. 基于 CNN 的手势识别[J]. 电脑编程技巧与维护, 2023(5): 104-106, 122.
- [9] 侯礼灏, 迟玉红, 谢虹铭. 基于知识图谱的法律问答系统研究[J]. 电脑编程技巧与维护, 2023(8): 106-108.
- [10] Devi, G. and Madhavi, D. (2024) Detection and Classification of Diabetic Retinopathy through Identification of Blood Vessel Thickness Using FOF& ML Classifiers. *Journal of Harbin Institute of Technology (New Series)*, 6, 84-96. <https://kns.cnki.net/kcms/detail/23.1378.T.20241018.1036.002.html>
- [11] Jose, A., Amrutha, S., Jain, J. and Joseena, S. (2021) A Study to Assess the Effect of Workshop on Knowledge Regarding Learning Disabilities among Nurses Working in Selected Institutions at Kottayam. *International Journal of Advances in Nursing Management*, 9, 46-48. <https://doi.org/10.5958/2454-2652.2021.00013.5>

-
- [12] Plc, F.G.S. (2018) SVM GLOBAL FUND PLC—ANNUAL REPORT—SEPTEMBER 30, 2017.
 - [13] Plc, F.G.S. (2017) SVM GLOBAL FUND PLC—ANNUAL REPORT—SEPTEMBER 30, 2016.
 - [14] Thomas, W. and Singh, V. (2024) Accelerating Neural Network Inference in Handwritten Digit Recognition—Comparative Study. *Parallel Processing Letters*, **34**, Article ID: 2430001. <https://doi.org/10.1142/s0129626424300010>
 - [15] Lin, Z., Li, J., Dai, G., Chen, T., Huang, S. and Lin, J. (2024) Contrastive Representation Enhancement and Learning for Handwritten Mathematical Expression Recognition. *Pattern Recognition Letters*, **186**, 14-20. <https://doi.org/10.1016/j.patrec.2024.08.021>
 - [16] Xu, H. and Zhang, X. (2024) Recognizing Digital Ink Chinese Characters Written by International Students Using a Residual Network with 1-Dimensional Dilated Convolution. *Information*, **15**, Article 531. <https://doi.org/10.3390/info15090531>
 - [17] Abdo, H.A., Abdu, A., Al-Antari, M.A., Manza, R.R., Talo, M., Gu, Y.H., *et al.* (2024) End-to-End Deep Learning Framework for Arabic Handwritten Legal Amount Recognition and Digital Courtesy Conversion. *Mathematics*, **12**, Article 2256. <https://doi.org/10.3390/math12142256>