IMAGE: A MAP OF THE STARS OF THE ORION CONSTELLATION

JournalPreview

London Journal of Research in Computer Science & Technology

Volume 24 | Issue 2 | Compilation 1.0

Great Britain Journals Press

WWW.JOURNALSPRESS.COM

JournalPreview

London Journal of Research in Computer Science & Technology

This document is a pre-published view of London Journal of Research in Computer Science & Technology Volume 24, Issue 2 and Compilation 1.0. For any minor changes and updations kindly follow your paper's live editing URL given in given in sent email or get in touch with our support team at support@journalspress.com or visit our website to use live chat support. This is a beta document thus order, content or existence of papers may alter in the published eJournal. You are requested to kindly acknowledge and approve your research paper in this JounalPreview within three days.

Journal Content

In this Issue

Scan to know paper details and author's profile

Practical Screening Method for Cancer Gene Diagnosis -How to Choose Cancer and Normal Patients bY four PrinciPles

Shuichi Shinmura

Emeritus of Seikei University

ABSTRACT

We developed a new theory of discriminant analysis (Theory1). Physicians can use it for practical medical diagnoses. Only Revised IP Optimal-LDF (RIP) obtains the minimum number of misclassification (MNM). RIP can discriminate linearly separable data (LSD) theoretically. It discriminated against 169 microarrays with two classes and found that 169 MNMs are zero and LSD. It can split high-dimensional arrays into many small LSD with less than n (patient's number) genes that are the candidates of multivariate oncogenes. We completed a new theory of high-dimensional gene data analysis (Theory2). A 100-fold Cross-Validation (Method1) can rank all candidates for the importance of diagnosis. Thus, if physicians firstly use Theory2 as the screening method, they can start their medical studies with the correct small sizes of candidates. This paper analyzes four arrays in detail and proposes correctly choosing cancer and normal patients using four principles.

Keywords: four universal data structures of 169 arrays, liver (gse14520, 357 patients), breast(gse42568, 116 patients), colorectal (gse8671, 63 patients), renal (gse66270, 28 patients).

Classification: LCC Code: RC268.57

Language: English

Great Britain Journals Press

LJP Copyright ID: 975814
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science & Technology

Volume 24 | Issue 2 | Compilation 1.0

Practical Screening Method for Cancer Gene Diagnosis -How to Choose Cancer and Normal Patients by four Principles

Shuichi Shinmura

ABSTRACT

We developed a new theory of discriminant analysis (Theory1). Physicians can use it for practical medical diagnoses. Only Revised IP Optimal-LDF (RIP) obtains the minimum number of misclassification (MNM). RIP can discriminate linearly separable data (LSD) theoretically. It discriminated against 169 microarrays with two classes and found that 169 MNMs are zero and LSD. It can split high-dimensional arrays into many small LSD with less than n (patient's number) genes that are the candidates of multivariate oncogenes. We completed a new theory of high-dimensional gene data analysis (Theory2). A 100-fold Cross-Validation (Method1) can rank all candidates for the importance of diagnosis. Thus, if physicians firstly use Theory2 as the screening method, they can start their medical studies with the correct small sizes of candidates. This paper analyzes four arrays in detail and proposes correctly choosing cancer and normal patients using four principles.

Keyword: four universal data structures of 169 arrays, liver (gse14520, 357 patients), breast(gse42568, 116 patients), colorectal (gse8671, 63 patients), renal (gse66270, 28 patients).

Author: Professor Emeritus of Seikei University, Economics.

I. INTRODUCTION

We completed a "New Discriminant Theory after R. A. Fisher (Theory1)" in 2015 [1-6]. After graduating from University, we worked at a private company and attended the project of the ECG diagnostic system at the Osaka Center for Cancer Cardiovascular Diseases (Center) and NEC. Dr. Nomura gave us ECG data over 1,000 patients having about 100 variables. These data (n≒1,000 * p≒100) consisted of one normal and over ten abnormal symptoms. He asked us to develop a diagnostic logic. From 1971 to 1974, we studied Fisher's linear discriminant function (LDF), quadratic discriminant function (QDF), and other statistical methods. We developed the diagnostic logic by LDF and QDF. Discriminant functions are weak for the discrimination of over three classes. We created several discriminant data with two classes, one normal symptom and one of the abnormal symptoms.

Our four-year study was inferior to Nomura's decision tree logic. This failure motivates us to develop the valuable discriminant theory for medical diagnosis. The number of misclassifications (NM) and error rate (ER) can evaluate discriminant results. However, those are unreliable because each discriminant functions have different NMs and ERs. Many defects of ER are the first problem of discriminant analysis (Problem1). We developed a Revised Optimal-LDF (RIP) that finds the minimum NM (MNM) that decreases monotonously ($MNM_k >= MNM_{(k+1)}$). This fact is Fact2 of discriminant analysis. If MNM=0, the data are linearly separable data (LSD). RIP by integer programming (IP) minimizes the NM and can find the MNM unique for the discriminant data. The first fact (Fact1) can explain the relation of LDF coefficients and NMs [1]. N linear equations made from n-cases divide the p-dimensional coefficients space into a finite convex polyhedron (CP). All interior points of each CP correspond infinite LDFs that have a unique NM=k and misclassify the same k-cases. There may be several optimal CPs (OCPs) having MNM. Only RIP can find one of the interior points of OCP theoretically. Therefore, RIP can quickly find 169 arrays are LSD. Because RIP can find the minimum ER, we need not compare many ERs of discriminant methods. Although many studies compare many ERs of classifiers, their efforts are useless.

RIP finds that Swiss banknote data consisting of 100 genuine and 100 counterfeit bills (200*6) [2, 6, 7] is LSD. A two variable model (x4, x6) is a minimum 2-dimensional LSD called a basic gene set (BGS) similar to Yamanaka's four genes. The 16 MNMs, including (x4, x6) among 63 models, are zero and LSD. The other 47 models are not LSD. We can explain this truth by Fact2. If MNM_k is MNM having k-variables, several $MNM_{(k+1)}$ are the MNMs added one variable to the former k-variables. In the cancer gene data analysis (Theory2), Fact2 explains array has the Matryoshka data structure that includes small Matryoshkas in it up to BGS (Structure1). Thus, the array has huge multivariate oncogenes candidates. If physicians study the characteristic of BGS by oncogenes, we expect they can find the new meaning of multivariate oncogenes (Validation4).

All discriminant functions, except for RIP and hard margin SVM (H-SVM) [8], cannot correctly discriminate LSD theoretically. However, H-SVM cannot discriminate overlapping data by computational error. MNM=0 means the first statistic for LSD. We check all NMs of SMs and BGSs by eight discriminant functions. Because all NMs of logistic regression is zero, only logistic regression using the maximum-likelihood algorithm can discriminate LSD empirically. Except for three LDFs, other discriminant functions based on the variance-covariance matrices cannot discriminate LSD correctly. Because of no study of LSD discrimination (Problem2), most researchers do not know the exact meaning of LSD. Thus, all studies are wrong and useless for cancer gene diagnosis.

We realized all pass/fail determination of exams is naturally LSD [6]. However, some ERs of exams are over 30% by Fisher's LDF. QDF misclassifies all pass students into the failed class if all pass students correctly answer items. This truth is the defect of the generalized inverse matrix algorism (Problem3). The discriminant theory is not inferential statistics (Problem4) because of no standard errors (SE) about discriminant coefficients and ER. We developed the 100-fold cross-validation (Method1) to estimate the experimental SE of both discriminant coefficients and ER. Many discriminant users validate their discriminant results by the leave-one-out method [9] or its extended k-fold CV. Both methods ignore the statistical principle of the relation of population and samples. Method1 is very easy and valuable. It copies the original data 100 times. That is the test sample and becomes the pseudo population. We add a random number to the test sample and sort it by the random values. We divide the sorted data into 100 training samples. Each training sample is a sample from the unique test sample. We call the averages of 100 ERs of training and test samples M1 and M2. M2=0 means that the original data and 100 test samples are LSD. We think SMs and BGSs with M2=0 are helpful for cancer diagnosis (Validation1).

After 1995, six medical projects studied the array profilings and released six first-generation arrays (old arrays) [10-15]. Many statisticians and engineers studied high-dimensional data analysis using these arrays. Because we completed Theory1 in 2015, we analyzed six old arrays downloaded from Higgins HP [16] as an applied problem of Theory1. RIP discriminated against six arrays and found those are LSD. Moreover, RIP could split six arrays into many Small Matryoshkas (SMs) within 54 days. Therefore, we completed the fundamental theory of gene data analysis (Theory2) until 2018 [17]. Until 2020, we confirmed six array results by 163 second-generation arrays (new arrays).

Because Theory1 and Theory2 are effortless and powerful, we expect medical specialists use those as a screening method at the entry point of medical diagnosis. Especially, SMs and BGSs with M2=0 become a beacon of medical research. This paper analyzes four arrays among 163 new arrays in detail and proposes the design principle of the array. Almost all analyses will finish in about one week.

II. MATERIAL AND METHOD

We explain materials and theories in three stages. Stages 1 introduces the summary of six old data, and Stages 2 introduces the summary of 73 arrays with two classes. Stages 3 introduces the new detailed analysis of four arrays among 73 arrays.

2.1 Six Old Arrays

Six medical projects have studied expression profiling of old arrays and published their papers. October 28th, 2015, we discriminated against Shipp data (77*7129) by RIP. RIP finds the MNM is zero, and Shipp array is LSD within one second.

Program1 is a program of the basic RIP coded by Mathematical Programming (MP) software named LINGO [18]. In [6], we explain RIP and Program1 in detail.

Moreover, we found that only 32 coefficients were not zero, and the other 7,097 coefficients were zero. We can naturally select 32 coefficients. Thus, other feature selection methods are useless. We can explain the reason for the surprising result by Fact1. Because Fact1 tells us the domain of RIP is less than n dimensions, RIP can find less than n=77 genes. We need to understand that the range of NM is [0, n] having (n+1) integer values.

We call the 32 genes the first SM (SM1). In Theory2, we call LSD Matryoshka. We discriminated against 7,097 genes again and found the second SM2. We developed a Matryoshka feature selection method (Method2) and coded Program3 by LINGO explained in [17]. Program3 can split Shipp's array into 222 Type1 SMs (MNM=0) and one Type2 SM (MNM >= 1). Other studies find only one gene set, MNM of which is over 1. Those gene sets are not multivariate oncogenes

Table1shows the summary of six arrays. The range of genes included in Type1 222 SMs is [18, 62]. MNM of Type2 SM is one and not LSD. JMP's logistic regression Degree of Freedoms (DF) range is [76, 80]. For the array (n<<p), DF generally becomes almost the same patients' numbers n=77 instead of p. Thus, logistic regression can split the array into many DF gene sets with almost n genes (Structure4). All studies ignored this linear model principle restricted by the simultaneous solution condition. Only Golub arrays have SMs of 170 Type1 and 22 Type2. Type2's MNM range is [1, 24]. Most arrays have many Type1 SMs and one Type2 SM. No studies have found four data structures of LSD until now.

Table 1: Two Types of SM and DF Decomposition

	n*p	Type1	Genes	Range	Type2 (Genes)	MNM	DF
Alon [10]	62*2000	64	1,999	[21, 42]	One SM with 1 gene	25	[61, 64]
Singh [14]	102*12625	215	12,598	[34, 85]	One SM with 27 genes	25	[85, 108]
Shipp [13]	77*7129	222	7,085	**[18, 62]**	One SM with 44 genes	**1**	**[76, 80]**
Tien [15]	73*12625	101	12,566	[100,139]	One SM with 59 genes	4	[172,180]
Chiaretti [11]	128*12625	155	12,623	[32, 122]	One SM with 2 genes	47	[99, 134]
Golub [12]	72*7129	**170**	6,348	[16, 56]	**22** SMs with 359 genes	**[1,24]**	[70, 75]

Although many engineering researchers proposed feature selection methods (FSs), they found only one gene-set about 50 genes using one- or two-variables statistical methods. Their gene sets are not multivariate oncogenes and LSD. Furthermore, they ignored other genes as useless genes. Therefore, we concluded that all engineering studies have many mistakes and are useless for practical diagnosis [23]. Theory2 finds that array has four vital data structures:

1. Six arrays are LSD. Those have the Matryoshka data structure that includes much smaller LSD, up to BGS (Structure1).
2. We can split the array into many SMs (Structure2), BGSs (Structure3), and DF gene sets (Structure4).

Table 2 shows the validations of SMs by Method1. We summarize 170 SMs of Golub into five groups. MIN, MAX, and MEAN columns are the minimum, maximum, and average of 35 ERs of test samples. Because all NMs of training samples are zero, those are LSD. At first, we consider these results are promising. However, we find many test samples with M2=0 among SMs and BGSs of 163 new second-generation arrays registered on the GSE database after 2007. Thus, M2 can evaluate its important ranking of SMs and BGSs. Especially, SMs and BGSs with M2=0 are useful for diagnosis.

Table 2: Evaluation 170 SMs of Golub

	MIN	MAX	MEAN (M2)
SM1~SM35	[0, 4.17]	[4.17, 18.1]	[0.42, 9.17]
SM36~SM70	[0, 4.17]	[4.17,18.0]	[0.42, 9.17]
SM71~SM105	[0, 9.72]	[8.33,23.6]	[5, 14.72]
SM106~SM140	[1.39, 12.5]	[11.11, 25]	[7.08, 16.11]
SM141~SM170	[0, 13.89]	[4.17,27.78]	[0.42, 18.89]

Program4 split Alon's array into Type1 129 BGSs with 2,000 genes, and there are no Type2 BGSs. Table 3 shows the BGS3's results among 129 BGS by Method1. It consists of 22 normal subjects and 40 tumor patients. Although all 100 MNMs of training samples are zero, 100 test samples have three different results. NMs of the first samples are 4,000, which means that all cancers are misclassified into the normal. NMs of the 100th sample are 2,200, which means that all normal are misclassified into the tumors. Method1 is helpful in the multivariate candidate's evaluation, also. Because Alon selected 2,000 from 6,500 genes, we consider their FS to omit several essential genes.

Table 3: Typical third BGS3 results among Alon's 129 BGS by 100-fold CV

		...37	...100	MIN	MAX	MEAN
Training Samples	0	0	0	0	0	0
Test Samples	4000	900	2200	14.52	64.52	45.44

Program3 and 4 split Tien's array into 101 SMs and 561 BGSs, respectively. Physicians had found over 100 driver oncogenes and tumor suppressor genes. We count 100 oncogenes included in 101 SMs. One SM includes three oncogenes. Physicians can understand the meaning of this SM by three genes. Most SMs do not include oncogenesis. We propose this simple check as Validation4. Golub, Shipp, and Singh belonged at the same group of Harvard Institute. They used the weighted voting system for their original FS and selected about 50 genes included in the neighborhood of over k oncogenes. Although they consider the relation of their selected genes and the oncogenes, other researchers ignored these legacy oncogenes. We consider four validations of SMs and BGSs. Validation1 is M2 by Method1, and Validation2 is RatioSV. Validation3 is the check of Ward and PCA. Validation4 is the check of the legacy oncogenes included in BGSs. Validation5 uses the genome cohort.

We discriminated SMs and BGSs by eight LDFs and QDF. Those are six MP-based LDFs in addition to logistic regression and Fisher's LDF. SVM4 and SVM1 are the soft-margin SVM using the penalty c=10,000 and 1. Because the quadratic programming (QP) defines H-SVM, SVM4, and SVM1, non-zero coefficients are few. However, no studies have been conducted on why QP-LDFs are useless for FS. Table 4 shows the number of non-zero coefficients by six MP-based LDF.

In 2018, we completed the basic Theory2 by six old arrays. We developed Program3 and Program4, and RatioSV. Because RIP discriminant scores (RipDSs) show the malignancy indexes for diagnosis, we make a signal data having all RipDSs instead of genes. JMP [19], the statistical software, analyzes signal data instead of the high-dimensional array. We propose several medical research themes in [17].

Table 4: The number of non-zero coefficients by six MP-based LDF

	Alon (62*2,000)	Golub (72*7,129)	Shipp (77*7,129)
RIP	56	37	43
IPLP	34	19	31
LP	34	19	31
H-SVM	2000	6274	7124
SVM4	2000	7129	7129
SVM1	2000	6262	7123

2.2 The 73 additional arrays registered on CuMiDa Gene Database

Bruno et al. [20] developed a CuMiDa gene database that includes 78 arrays of 13 carcinomas registered on GSE gene DB after 2007. They selected these second-generation arrays after their quality check. It consists of five one-class arrays, 57 two-class arrays, and 16 with three to seven classes. We omit five one-class arrays from our analysis. We made 106 two-class arrays by two-class combinations from 16 data. Because we confirmed four data structures by 57 and 106 arrays, we completed Theory 2 [21-24] in 2022.

Table 5 shows the summary of 73 original arrays with over two classes. We show one or two arrays among 13 carcinomas. Type column shows the 13 carcinomas. GSE shows the GSE code. The n shows the number of cases. The case range is [12, 357] and 18 arrays have less than 30 cases. GENE shows Gene's number. Its range is [12621, 54676]. Because we wish to obtain the results quickly, we analyze 1,2621 genes, which are the maximum genes of six old arrays. R1 is the ratio of 1,2621 genes and all genes included in each array. Because Pancreatic and the sixth Brest6 contain 54,676 genes, two R1s are 0.23. Class means the 73 arrays' number of classes. We evaluated the first 10 SMs by Method1(100-fold CV). MinM2 shows the minimum values of M2 of first 10 SMs. The range is [0, 12.24]. Because six old arrays have no SMs with M2=0, we consider the new arrays' quality of expression becomes better than old arrays. Otherwise, new arrays include more useful new genes for diagnosis. Nation shows the nation of the first author. LSD1 and LSD2 are the numbers of all SMs and BGSs with M2=0(10-fold CV). At first, we evaluate 10 SMs by 100-fold CV. Because the check of 100-fold CV is bothered, we evaluate all SMs and BGSs by the 10-fold CV. We use Table 5 for the becon of 169 arrays's analyses.

Table 5: Results by CuMiDa's 169 Arrays and Theory2

ID	TYPE	GSE	n	GENE	R1	CLASS	MinM2	SVM	MLP	DT	NB	RF	Nation	LSD1	LSD2
	Max	MAX	357	54676	1	7	12.24	1	1	1	1	1	0	5	31
	Min	MIN	12	12621	0.231	2	0	0.26	0.29	0.25	0.34	0.34	0	0	0
	Total	SUM	5427	2867932		179	103.69	64.16	60.97	55.6	61.15	62.24	0	15	185
1	Pancreatic	GSE16515	51	54676	0.23	2	0.45	0.86	0.78	0.78	0.84	0.82	USA	0	0
7	Breast6	GSE42568	116	54676	0.23	2	0	0.99	0.99	0.94	0.99	0.97	Ireland	0	2
17	Liver3	GSE14520	357	22278	0.57	2	0.11	0.97	0.8	0.92	0.96	0.96	USA	0	0
24	Liver10	GSE14520	41	22278	0.57	2	0	1	0.98	0.98	0.95	1	USA	0	7
25	Throat1	GSE42743	103	54676	0.23	2	0	0.87	0.83	0.85	0.86	0.87	USA	0	1
34	Leukemia6	GSE28497	281	22284	0.57	7		0.88	0.72	0.73	0.78	0.79	USA		
46	Prostate9	GSE6919	124	12626	1.00	2	9.76	0.67	0.65	0.45	0.63	0.69	USA	0	0
51	Ovary4	GSE6008	98	22284	0.57	4		0.71	0.64	0.65	0.68	0.71	USA		
52	Brain1	GSE50161	130	54676	0.23	5	0.02	0.95	0.82	0.85	0.85	0.91	USA	0	0
54	Bladder1	GSE31189	85	54676	0.23	2	12.24	0.64	0.58	0.54	0.46	0.55	USA	0	0
58	Lung3	GSE19804	114	54676	0.23	2	1.75	0.93	0.85	0.91	0.91	0.92	Taiwan	0	0
62	Renal1	GSE66270	28	54676	0.23	2	0	1	1	0.79	1	1	Germany	5	31
68	Gastric3	GSE19826	24	54676	0.23	2	0.04	0.67	0.67	0.67	0.71	0.67	China	0	0
74	Colorectal6	GSE8671	63	54676	0.23	2	0	1	1	0.94	1	1	Swiss	4	16
77	Colorectal 9	GSE21510	147	54676	0.23	3		0.99	1	0.9	0.97	0.94	Japan		

Cilia et al. [25] analyze Golub's and Alon's arrays by Weka. Their FSs select several gene sets. They evaluate ERs of gene sets by several classifiers of Weka and 10-fold CV. Their main minimum ERs of Golub's and Alon's ERs are roughly 8%.

Bruno et al. analyze original 78 arrays by the Weka program. They calculate eight classifiers' accuracy rates (1 - ERs) of all arrays. SVM, RF, Multilayer Perceptron, DT, Naïve Bayes are the supervised learning methods to analyze 73 supervised learning data. Because k-nearest neighbors, k-means, and Hierarchical Clustering are the unsupervised learning methods and are proper to analyze five one-class unsupervised learning data, we omit three ones from Table 5.

Because many engineering researchers use Weka, we can understand their results' overview by this table. Two (Mean ± SD) of SVM and RF are (0.88 ± 0.14) and (0.85 ± 0.15). They recommend SVM and RF because both accuracy rates are better than others. Because the accuracy rates of Breast1 by SVM is 1, NM of SVM by three-fold CV is 0. If Weka's SVM is H-SVM, Breast1 is LSD. Because it is kernel-SVM, we cannot judge Breast1 is LSD or not. Thus, all classifiers cannot judge arrays are LSD or not because those discriminant hyperplanes are not linear. Only 16 arrays of 73 data can separate two-class by SVM. Because RIP finds all 163 arrays are LSD by 10-fold CV, we conclude that eight classifiers are useless for gene diagnosis.

Bruno's ER survey gives us a vital hint for the four categories of classifiers from the viewpoints of the reliability of ER or LSD discrimination.

First category: Only RIP, H-SVM, and logistic regression can correctly discriminate LSD.

Second category: NMs of statistical LDFs based on the variance-covariance matrices are unreliable, especially for LSD discrimination. Mostly, those are not zero. Six ERs of Chiaretti, Shipp, Alon, Singh, Golub, and Tien by Fisher's LDF using the singular value decompoition are 1%, 4%, 8%, 10%, 11%, and 17%, respectively.

Third category: QDF and kernel-SVM are non-linear discriminant functions. Even if NM is zero, it does not mean LSD. Fourth category: Because other classifiers such as the decision tree use the complex discriminant hyperplane, NM=0 does not mean LSD. We are afraid most researchers do not understand the correct definition of LSD. Although no researchers found LSD, we found many LSD. Those are Swiss banknote data, every exam data, Japanese 29 regular and 15 small cars, Fisher's Iris data [31], and 169 arrays.

Engineering researchers compare ERs of many classifiers and decide the best classifier for each data. Because RIP finds MNM and the minimum ER, we need not compare and select classifiers for LSD discrimination because RIP is the best theoretically [31]. If M2=0, it means that the original data, SMs and BGSs, training samples, and test samples are LSD. Now, MP-based LDFs, RIP and H-SVM, can discriminate LSD theoretically. JMP's logistic regression can discriminate LSD experimentally confirmed by over a thousand checks of SMs and BGSs.

Because we confirm four universal data structures by 169 arrays, we completed Theory2 in 2022.

2.3 Detail Study of Four Data

This paper analyzes four new arrays, Liver3 [26], Breast6 [27], Colorectal 6 [28], Renal 1 [29], precisely. We studied epidemiological data with Dr. Takaichirou Suzuki and others at the Center. We published many cancer diagnosis papers until 1983 [30]. Thus, we established three principles of patient design as follows:

1) Adequate sample size. Usually, statisticians think over 100 cases.

2) Equally sample sizes of both classes.

3) First analysis is to compare cancer with a control group. Physicians must analyze two different cancers after the first analysis. Although many arrays of 169 arrays ignore the above principles: 1) lack the normal class, 2) small sample size of fewer than 30 cases, 3) the unbalanced sample sizes.

These arrays do not obtain good data analysis results. We are suspicious that these projects do not expect vital gene data analysis results. However, we realize the fourth principle instead of the first principle by Liver3. Three projects, except for Colorectal6, have better modified their data by our advice and analyzed the modified array again. They will obtain many useful multivariate oncogenes within one week.

III. RESULT

3.1 Liver3

Liver3 consists of 181 HCC patients and 176 normal subjects. We expect a good result because it satisfies three principles: 1) Because the sample size is 357, it is enough data from a statistical viewpoint. Most statisticians expect over 100 cases. 2) The sizes of two-class are almost the same. 3) It has the normal control class. Moreover, it has the genome cohort for validation. However, we find the correct fourth principle only for gene expression. We consider Liver3 includes many specific patients interested in medical study. If physicians take those patients from the array and treat them as test samples, they can get better results by the medium sample. Or, they cangorize Liver 3 into several arrays.

Program3 split 12,671 genes into 89 SMs with 12,611 genes and SM90 with 60 genes (MNM=2) by 20 iterations of discrimination. CPU time is 2h13s. If Program3 discriminates twice, we obtain 40 SM with 12,615 by 8m22s. Because 40 SMs have many genes, we explain 89 SMs.

Table 6 shows 89 SMs' results by Method1. SM column shows the first 10 SMs. We omit other 79 results. Gene shows the gene's number of 10 SMs. The following three columns are the minimum, maximum, and average of 10 NMs for the test samples. The last three columns are the minimum, maximum, and averages of 10 ERs of test samples. However, last three rows are the summary of 89 SMs. M2 range is [0.18, 13.7] and the average is 5.1. There is no SMs with M2=0.

Table 6: 89 SMs by Method1

SM	gene	min	max	mean	MIN	MAX	MEAN
1	13	110	310	180	3.08	8.68	5.04
2	18	50	320	181	1.4	8.96	5.07
3	16	90	240	164	2.52	6.72	4.59
4	13	80	340	172	2.24	9.52	4.82
5	15	70	270	176	1.96	7.56	4.93
6	8	100	220	152	2.8	6.16	4.26
7	15	130	270	185	3.64	7.56	5.18
8	23	100	290	186	2.8	8.12	5.21
9	16	90	320	176	2.52	8.96	4.93
10	22	80	220	137	2.24	6.16	3.84
MIN	6	0	41	6.5	0	1.15	0.18
MAX	25	400	600	488	11.2	16.8	13.7
Mean	15	108	270.8	182.2	3.03	7.59	5.1

JMP uses two symbols and colors to identify cancer patients (×, Red) and normal subjects (○, Blue). Figure 1 shows Ward's two-way clustering (Left) and PCA (Right) of SM1. After the Ward analyzes SM1, PCA input the clustering color information. We strongly recommend these combinations analyses of both methods in Euclidian space. PCA has three plots. The middle scatter plot shows the most cancer patients on the first and fourth quadrants and the most normal subjects on the second and third quadrants. However, several cases overlap with another class. The cumulative contribution ratio of the first and second components is 23.73 (=16.5+7.23) %. It explains about 24% variation out of 357 cases variation. The left plot shows the eigenvalues, and the right plot shows the factor loading plot. Ward has three parts. The left part shows the 357 cases. The middle part is the two-way clustering. The right part is the case dendrogram. The bottom dendrogram shows the variable dendrogram. Although the upper red part corresponds to cancer patients and the lower blue part corresponds to the normal subjects, several cases are misclassified to another class like PCA's scatter plot. We strongly recommend taking off misclassified cases that may be specific cases from the analysis and treating those cases as test samples.

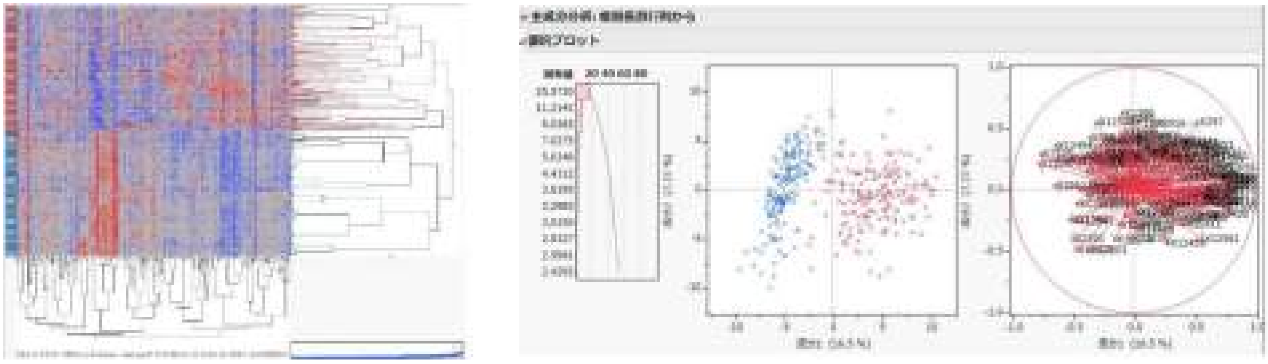


Figure 1: Ward cluster (Left figure) and PCA (Right figure) of SM1 of Liver (GSE14568)

Figure 2 shows SM89's 357 cases located in 138-dimensional gene space. The scatter plot shows that both cases overlap near the origin. Two classes locate on four quadrants, and the cumulative contribution ratio is 32.9 %. Because all SMs are LSD, we know two classes are separable in the other 136-dimensional subspace. However, we consider that physicians do not use SM89 for diagnosis by Validation3 judgement.

We strongly recommend taking off these misclassified cases found in Figure 1. The omitted cases are helpful for the validation of Theory2's result as same as the gene cohort. If researchers analyze the modified arrays, they can obtain more precise results from SM1.

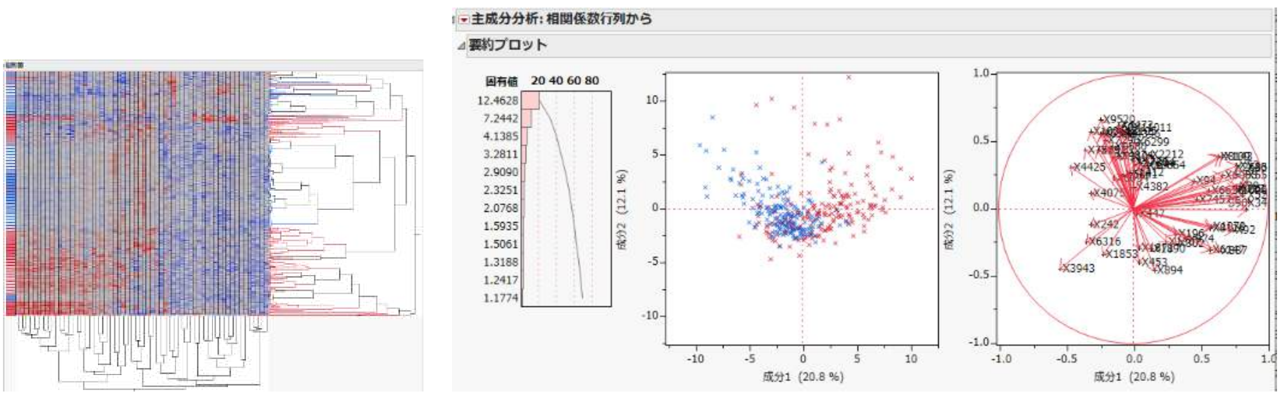


Figure 2: Ward cluster (Left figure) and PCA (Right figure) of SM89 of Liver (GSE14568)

We split 12,561 genes included in 89 SMs into BGSs by Program4. However, we stopped this work. We must calculate each BGS one by one, and it needs about two days. Because SM1's results are wrong, BGS brings no merit for diagnosis. If we confirm the excellent results of Method1 after Ward and PCA (Validation3), we had better obtain the BGS.

3.2 Breast6

Breast6 consists of 101 cancer patients and 15 normal subjects. Although the total case number is 116, we can forecast the bad results because the normal cases are few. If this project collects more than 35 normal subjects and selects 50 typical cancer patients, they can get marvelous results by 100 cases.

Program3 split 12,671 genes into 113 SMs with 12,618 genes and SM114 with 53 genes (MNM=1) by twice discrimination. CPU time is 6m29s. If Program3 discriminates 20 times, CPU time is 3h42m 16s. We obtain 359 SMs with 12,646 genes. Table 7 shows the results of 114 SMs by Method1. SM column shows only 13 SMs and SM114. The min, max, mean show three values of MNM of ten test samples. MIN, MAX, and MEAN are three values of ERs of ten test samples. There is no SMs with M2=0. The last three rows show the summary of 114 SMs. The range of 114 minimum, maximum and average values are [0, 7.8], [7.8, 21.55], and [4, 15], respectively. Results are bad because of the unbalanced data.

Table 7: 89 SMs and Other90 by Method1

SM	min	max	mean	MIN	MAX	MEAN
1	30	230	93	2.688	20.61	8.333
2	10	210	89	0.896	18.82	7.975
3	20	90	48	1.792	8.065	4.301
4	10	100	45	0.896	8.961	4.032
5	40	180	99	3.584	16.13	8.871
6	30	220	114	2.688	19.71	10.22
7	30	170	102	2.688	15.23	9.14
8	30	190	82	2.688	17.03	7.348
9	40	160	94	3.584	14.34	8.423
10	30	130	85	2.688	11.65	7.616
111	90	250	174	7.759	21.55	15
112	30	140	79	2.586	12.07	6.81
113	30	140	82	2.586	12.07	7.069
114	10	140	60	0.862	12.07	5.172
MIN	0	90	45	0	7.759	4.032
MAX	90	250	174	7.759	21.55	15
Mean	20.53	143.8	73.74	1.778	12.44	6.382

Figures 3 and 4 show both figures of SM1 and SM113. Figure 4 shows 101 cancer patients gathered around the origin. If we do not know the two classes' labels, we consider 15 normal subjects to be outliers of 101 patients. Normal subjects become the subset of cancer because of the small sample size. If the normal sample size is almost the same, we consider both classes more separable (Validation3). However, because these are our hypotheses, we must need further study.

1. Most researchers misunderstand the class information. Only discriminant functions can analyze the class information. Other statistical methods can analyze only the measured variables, not classified information, and those use the class information as a label.
2. Figures 3 and 4 show three clustering pairs and fourteen clustering pairs. According to clustering results, two classes of the scatter plot are almost overlapping.
3. We consider the unbalanced sizes to cause this problem.

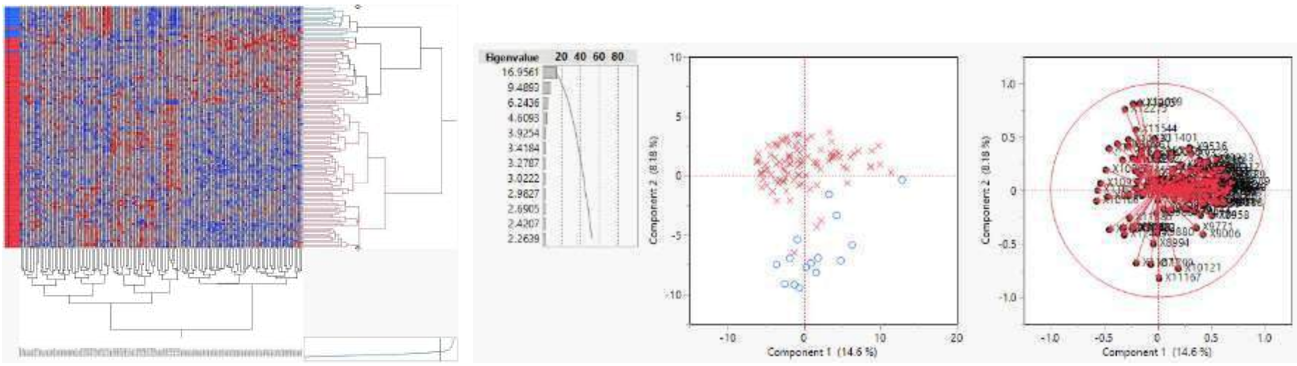


Figure 3: The two-way clustering and PCA of SM1 of Breast (GSE42568)

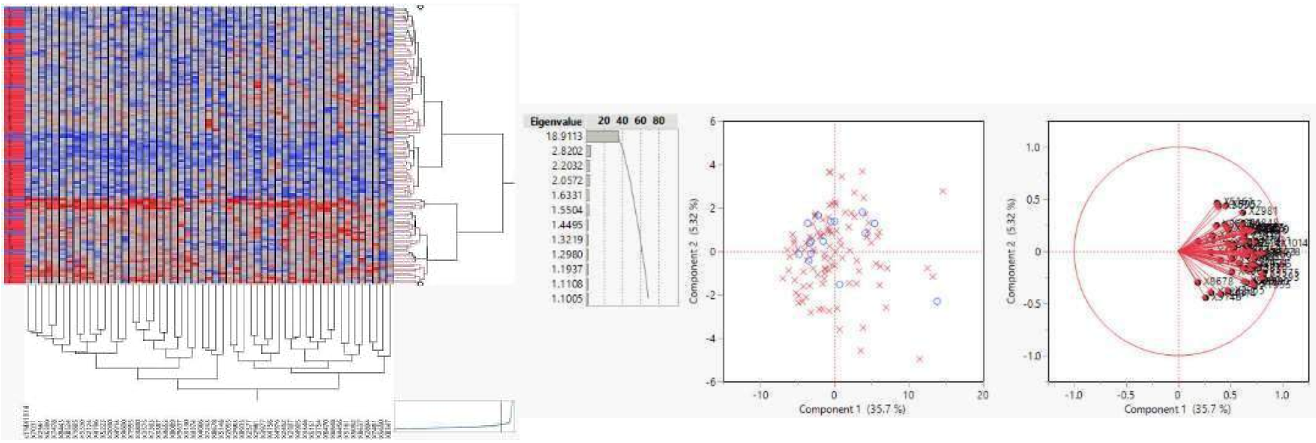


Figure 4: The two-way clustering and PCA of SM113 of Breast (GSE42568)

3.3 Colorectal6

Colorectal6 consists of 31 adenoma patients and 32 normal subjects. We doubt a good result because of less than 100 cases. Program3 split the array into 543 SMs with 12,609 genes and SM544 with 12 genes (MNM=3). Although Method1 evaluates 544 SMs, Table 8 shows 14 SMs with M2=0 and the other five SMs. The last column is RatioSV (Validation2). Because the result is better than Liver3 and Breast6, we consider the medium sample size is better than over 100 cases.

Table 8: 89 SMs and SM90 by Method1

SM	gene	min	max	mean	MIN	MAX	MEAN	min	max	mean	MIN	MAX	MEAN	RatioSV
1	13	0	0	0	0	0	0	0	0	0	0	0	0	37.368
2	18	0	0	0	0	0	0	0	30	10	0	4.7619	1.5873	44.147
3	16	0	0	0	0	0	0	0	0	0	0	0	0	52.927
17	13	0	0	0	0	0	0	0	0	0	0	0	0	41.707
20	11	0	0	0	0	0	0	0	0	0	0	0	0	39.023
23	6	0	0	0	0	0	0	0	0	0	0	0	0	23.022
24	15	0	0	0	0	0	0	0	0	0	0	0	0	39.078
40	9	0	0	0	0	0	0	0	0	0	0	0	0	32.497
45	11	0	0	0	0	0	0	0	0	0	0	0	0	26.214
50	11	0	0	0	0	0	0	0	0	0	0	0	0	16.627

53	12	0	0	0	0	0	0	0	0	0	0	0	0	20.968
54	17	0	0	0	0	0	0	0	0	0	0	0	0	39.879
106	18	0	0	0	0	0	0	0	0	0	0	0	0	32.224
141	11	0	0	0	0	0	0	0	0	0	0	0	0	24.246
165	14	0	0	0	0	0	0	0	0	0	0	0	0	28.344
541	45	0	0	0	0	0	0	110	160	134	17.46	25.397	21.27	-
542	38	0	0	0	0	0	0	80	170	131	12.698	26.984	20.794	-
543	41	0	0	0	0	0	0	100	140	121	15.873	22.222	19.206	-
544	12	0	2	0.7	0	3.17	1.11	30	150	77	4.7619	23.81	12.222	-
MIN	6	0	0	0	0	0	0	0	0	0	0	0	0	-
MAX	48	0	2	1.3	0	3.17	2.06	110	200	134	17.46	31.746	21.27	-
Mean	23.2	0	0.05	0.02	0	0.03	0.01	15.8	65.3	36.7	2.5152	10.361	5.8252	-

Figure 5 shows SM1's figures. Ward cluster shows two classes complete become two clusters. The scatter plot shows all normal subjects on the first and fourth quadrants and all cancer patients on the second and third quadrants. The cumulative contribution ratio is 57.2 (=44.1+13.1) %. Method1 finds many SMs with M2=0 (Validation1). RatioSVs have many large values (Validation2). Ward and PCA can separate two groups (Validation3). We are surprised these marvelous results.

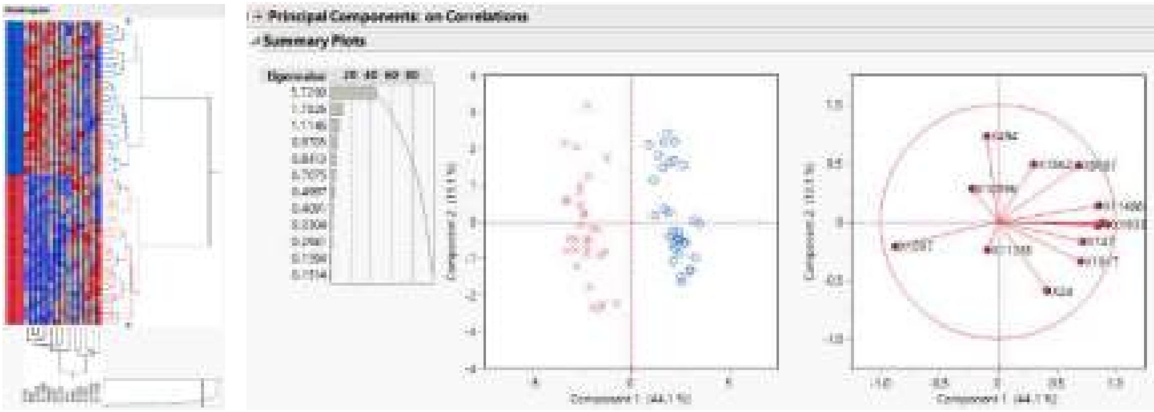


Figure 5: Ward cluster (Left figure) and PCA of SM1 of Colorectal (GSE8671)

Figure 6 shows SM543's figures. Ward cluster shows two classes become 18 small clusters. The scatter plot shows two classes overlap. Although the cumulative contribution ratio is 40.9 (=30.8+10.1) %, the scatter plot cannot show LSD. Although two classes become LSD on other 10-dimensional space, it is useless for diagnosis. Validation3 is useful for selecting vital multivariate oncogenes.

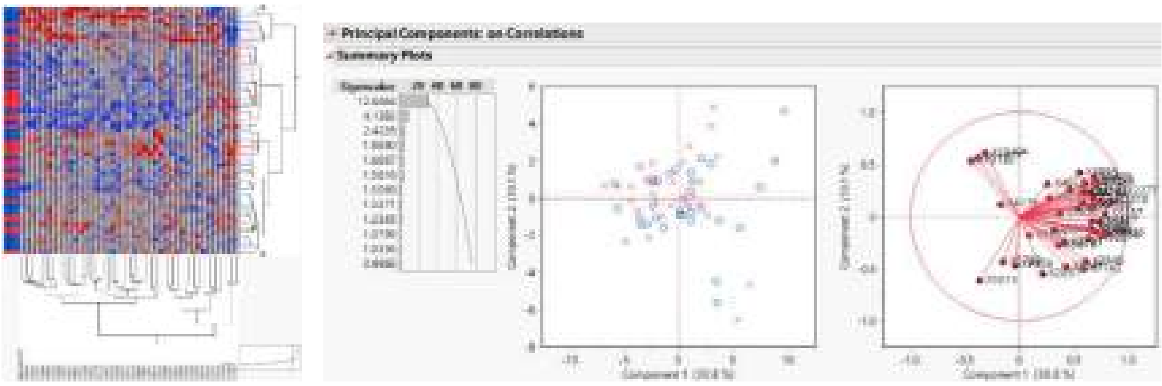


Figure 6: Ward cluster (Left figure) and PCA of SM89 of Colorectal (GSE8671)

3.4 Renal1

We never recommend the following analysis for arrays having less than 30 cases. If researchers agree with our claim, they collect about an additional 30 patients and analyze the same with Colorectal6. Renal1 consists of 14 normal subjects and 14 cancer patients. Although Renal1 does not satisfy Principle1, we misunderstood that the analysis of Renal1 is easy because of the small sample at first. Thus, we analyzed it step by step by Theory2 in detail. However, we spent two weeks and got awful results.

Program3 quickly splits the first 12,621 genes into 1,066 SMs (MNM=0) and the last SM1067 with one gene (MNM=14). Here, we had better stop the analysis and estimate the final SMs as follows. Because the average of genes included in 1066 SMs is 11.83 and Renal1 has 54,676 genes, it may consist of about 4,618 SMs. We had better estimate the number of SMs.

However, we skipped the above estimate. Next, Program4 split the genes included in 1,066 SMs into BGSs. After obtaining both results, we evaluated both SMs and BGSs by Method1. Physicians never follow this procedure, especially for a small sample. Physicians analyze case-by-case. For small samples, we recommend the following steps. SM decomposition is easy. Next, Program4 split the genes contained in SM1 into BGS, as shown in Table 9. Then, Method1 evaluates SM1 and 8 BGSs. We consider other small samples' results may be wrong as same as the table. This analysis' results show two critical facts. The numbers of genes of 6 BGSs are one, and the other two BGSs are 3. This result means that six single-genes can easily discriminate 28 patients into six LSDs.

Table 9: The first 10 SMs with 38 BGSs and 10 SMs with MNM >=1

RIP	Gene	min	max	mean	MIN	MAX	MEAN
SM1	12	1400	1400	1400	50	50	50
BGS11	1	1400	1400	1400	50	50	50
BGS12	1	1400	1400	1400	50	50	50
BGS13	1	1400	1400	1400	50	50	50
BGS14	1	1400	1400	1400	50	50	50
BGS15	1	1400	1400	1400	50	50	50
BGS16	3	1400	1400	1400	50	50	50
BGS17	1	1400	1400	1400	50	50	50
BGS18	3	1400	1400	1400	50	50	50

However, Method1 tells us that NMs and ERs of 100 test samples are very bad. As shown by the minimum, maximum, and average values, 1400 cases, which is half of the 2,800 cases, are misclassified, and all Minimum, Maximum, and Average ERs are 50%. M2=50 becomes the stopping rule of M2 for small data. Although most researchers use a k-fold CV, they cannot obtain the same severe result as Method1. Fig. 7 shows two figures by Ward and PCA of SM1. Because the results are so good, this project have better added cancer and normal patients by the same criteria. If the results are wrong, they must change the patient selection criterion.

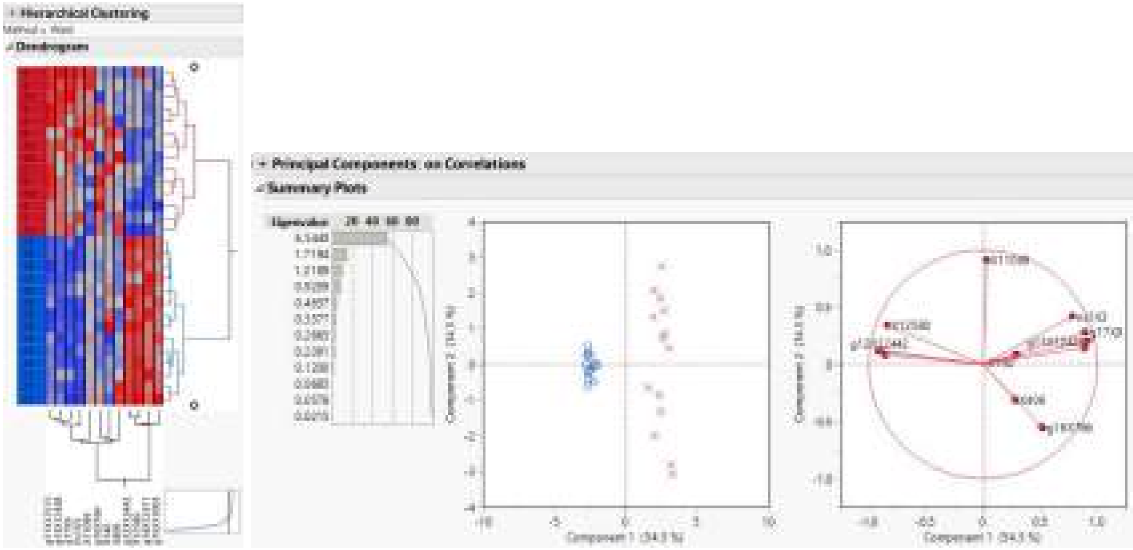


Figure 7: Both the figures of the Ward and PCA about SM1 of Renal (GSE66270)

It is important that only Colorectal6 has good results in Validation1 and Validation3. Now, we cannot decide the proper threshold of RatioSV (Validation2).

IV. DISCUSSION

Theory1 solved four discriminant problems by RIP and Method1. NM and ER evaluate discriminant results. Thus, we developed RIP that finds MNM. Because discriminant analysis is not inferential statistics and ER is the most vital statistics defined by class information, we found two vital facts of discriminant analysis. Fact1 tells us the relation of LDF coefficients and NM. Only RIP finds MNM that decreases monotonously (Fact2). Therefore, RIP can easily find LSD and confirms that 169 arrays are LSD. RIP can split the array into many SMs and BGSs. At last, we find Fact3 that 169 arrays have four universal data structures similar to many Matryoshka nested dolls [31].

This paper analyzes four arrays, such as Liver3, Breast6, Colorectal6, and Renal1. We consider Liver3 includes several specific patients important from the medical study. In general, we expect the best results of Liver3 that satisfy three design principles of cancer and normal subjects. However, Validation1 (Method1) and two figures (Validation3) tell us the problems.

Program3 and 4 find many SMs and BGSs candidates of multivariate oncogenes. Method1 (Validation1) gives us the vital ranks of diagnosis, especially M2=0. Because our screening methods by Theory1 and Theory2 almost finish within one week, every physician analyzes their DNA and RNA gene data as the screening methods at the first study before medical research. They can save much research time and obtain many correct multivariate oncogenes LSD. We completed the discriminant theory of LSD useful for all discriminant data [31] in 2023.

We sincerely hope physicians will achieve to overcome cancer by our methods.

ACKNOWLEDGMENT

Our research depends on the powerful LINGO solver supported by L. Schrage, K. Cunningham, and H. Ichikawa (LINDO Japan). We could establish Theory2 by IP solver and statistical software using suitable arrays. J. Sall developed JMP, a powerful statistical package. Bruno et al. developed CuMiDa. Jeffery et al. uploaded six old data. Thanks also for the suppliers of 73 medical studies. The late father's legacy, Otojirou, was helpful in these studies.

REFERENCE

1. Shinmura S (2000) A new algorithm of the linear discriminant function using integer programming. New Trends in Probability and Statistics, 5: 133-142.
2. Shinmura S (2010) The optimal linear discriminant function. Union of Japanese Scientist and Engineer Publishing, Japan (ISBN 978-4-8171-9364-3).
3. Shinmura S (2011) Problems of discriminant analysis by mark sense test data. Japanese Soc Appl Stat 4012:157-172.
4. Shinmura S (2014) End of discriminant functions based on variance-covariance matrices. ICORE2014: 5-16.
5. Shinmura S (2015) Four serious problems and new facts of the discriminant analysis. In: Pinson E et al. (ed) Operations research and enterprise systems. Springer, Berlin: 15-30
6. Shinmura S (2016) New Theory of Discriminant Analysis After R. Fisher. Springer.
7. Flury B, Riedwyl H (1988) Multivariate statistics: a practical approach. Cambridge University Press, New York.
8. Vapnik V (1999) The Nature of Statistical Learning Theory. Springer.
9. Lachenbruch PA, Mickey MR (1968) Estimation of error rates in the discriminant analysis. Technometrics. 10 (1): 11.
10. Alon U et al. (1999) Broad Patterns of Gene Expression Revealed by Clustering Analysis of cancer and Normal Colon Tissues Probed by Oligonucleotide Arrays. Proc. Natl. Acad. Sci. USA, 96: 6745-6750.
11. Chiaretti S et al. (2004) Gene expression profile of adult T-cell acute lymphocytic leukemia identifies distinct subsets of patients with different responses to therapy and survival. Blood. April 01st, 2004, 103/7: 2771-2778.
12. Golub TR et al. (1999) Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. Science. 1999 October 15th, 286/5439: 531-537.
13. Shipp MA et al. (2002) Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. Nature Medicine 8: 68-74
14. Singh D et al. (2002) Gene expression correlates of clinical prostate cancer behavior. Cancer Cell: March 2002, Vol.1: 203- 209
15. Tian E et al. (2003) The Role of the Wnt-Signaling Antagonist DKK1 in the Development of Osteolytic Lesions in Multiple Myeloma. The New England Journal of Medicine, Vol. 349, 26: 2483-249
16. Jeffery IB, Higgins DG, Culhane C (2006) Comparison and evaluation of methods for generating differentially expressed gene lists from microarray data. BMC Bioinformatics:1-16
17. Shinmura S (2019a) High-dimensional Microarray Data Analysis. Springer, Dec.
18. Schrage L (2006) Optimization Modeling with LINGO. LINDO Systems Inc.
19. Sall JP, Creighton L, Lehman A (2004) JMP Start Statistics, Third Edition. SAS Institute Inc. (Shinmura, supervise Japanese version)
20. Bruno CF, Eduardo BC, Bruno IG, Marcio D (2019) CuMiDa: An Extensively Curated Microarray Database for Benchmarking and Testing of Machine Learning Approaches in Cancer Research. Journal of Computational Biology. 26-0: 1-11

21. Shinmura S (2019b) Release from the Curse of High Dimensional Data Analysis. Big Data, Cloud Computing, and Data Science Engineering (Studies in Computational Intelligence 844):173-196
22. Shinmura S (2020c) First Success of Cancer Gene Data Analysis of 169 Microarrays for Medical Diagnosis. CSCI-ISCB: COMPUTATIONAL BIOLOGY1-7. Transactions on Computational Science & Computational Intelligence, Springer Nature.
23. Shinmura S (2021a) Twenty-three Serious Mistakes of Cancer Gene Data Analysis since 1995. In: Arabnia HR et al. (eds.), Advances in Computer Vision and Computational Biology, Transaction on Computational Science and Computational Intelligence, https://doi.org /10.1007/973-3-030-71051-4_ 62. Springer Nature Switzerland AG 2021: 805-822 (in Press)
24. Shinmura S (2021b) First Theory of Cancer Gene Data Analysis of 169 Microarrays and Four Universal Data Structures for Big Data. CSCI-ISCB: COMPUTATIONAL BIOLOGY1-14. Transactions on Computational Science & Computational Intelligence, Springer Nature (in Press).
25. Cilia ND et al. (2019) An Experimental Comparison of Feature-Selection and Classification Methods for Microarray Datasets. Information 10, 109: 1-13
26. Roessler S, Jia HL, Budhu A, Forgues M, et al. A unique metastasis gene signature enables prediction of tumor relapse in early stage hepatocellular carcinoma patients. Cancer Res 2010 Dec 15;70(24):10202-12. PMID: 21159642
27. Clarke C, Madden SF, Doolan P, Aherne ST et al. Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. Carcinogenesis 2013 Oct;34(10):2300-8. PMID: 23740839
28. Sabates-Bellver J, Van der Flier LG, de Palo M, Cattaneo E et al. Transcriptome profile of human colorectal adenomas. Mol Cancer Res 2007 Dec;5(12):1263-75. PMID: 18171984
29. Wotschofsky Z, Gummlich L, Liep J, Stephan C et al. Integrated microRNA and mRNA Signature Associated with the Transition from the Locally Confined to the Metastasized Clear Cell Renal Cell Carcinoma Exemplified by miR-146-5p. PLoS One 2016;11(2):e0148746. PMID: 26859141
30. Shinmura S, Suzuki T, Koyama H, Nakanisshi K (1983) Standardization of medical data analysis using various discriminant methods on a theme of breast diseases, Medinfo 83, J.F. Van Bemmel and O Wigertz, pp.349-352, North-Holland Publishing Company.
31. Shinmura S (20124) The First Discriminant Theory of Linearly Separable Data. Springer.

Scan to know paper details and author's profile

Smooth Attention: Improving Image Semantic Segmentation

Kriuk Boris, Kriuk Fedor & Praveen Karthik*

Hong Kong University of Science and Technology

ABSTRACT

Attention mechanisms have become a fundamental component of deep learning, including the field of computer vision. The key idea behind attention in computer vision is to help the model focus on the relevant spatial regions of the input image, rather than treating all regions equally. The traditional approaches to attention mechanisms in computer vision often suffer from distribution inconsistencies in the attention maps, resulting in sharp transitions that negatively affect model's focus and lead to poor generalization on complex shapes. The problem of spatial incoherence is particularly pronounced in the task of semantic segmentation, where accurate pixel-level predictions require a detailed understanding of the spatial relationships within the image. In this paper, we propose an attention mechanism called Smooth Attention designed for convolutional neural networks to address the problem of spatial inconsistency in attention maps through multidimensional spatial smoothing. We conduct a series of experiments to evaluate the effectiveness of the proposed mechanism and demonstrate its superior performance compared to traditional methods.

Keywords: semantic segmentation, attention mechanism, spatial relationships, computer vision, deep learning.

Classification: MSC Code: 68T45

Language: English

Great Britain Journals Press

LJP Copyright ID: 975815
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science & Technology

Volume 24 | Issue 2 | Compilation 1.0

Smooth Attention: Improving Image Semantic Segmentation

Kriuk Boris[α,*] Kriuk Fedor [σ] & Praveen Karthik[ρ]

ABSTRACT

Attention mechanisms have become a fundamental component of deep learning, including the field of computer vision. The key idea behind attention in computer vision is to help the model focus on the relevant spatial regions of the input image, rather than treating all regions equally. The traditional approaches to attention mechanisms in computer vision often suffer from distribution inconsistencies in the attention maps, resulting in sharp transitions that negatively affect model's focus and lead to poor generalization on complex shapes. The problem of spatial incoherence is particularly pronounced in the task of semantic segmentation, where accurate pixel-level predictions require a detailed understanding of the spatial relationships within the image. In this paper, we propose an attention mechanism called Smooth Attention designed for convolutional neural networks to address the problem of spatial inconsistency in attention maps through multidimensional spatial smoothing. We conduct a series of experiments to evaluate the effectiveness of the proposed mechanism and demonstrate its superior performance compared to traditional methods.

Keywords: semantic segmentation, attention mechanism, spatial relationships, computer vision, deep learning.

Author α σ ρ: Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong.
α σ ρ: Sparcus Technologies Limited, Suit C, 50 Stanley Street, Central, Hong Kong.

I. INTRODUCTION

Attention mechanisms[1–3] have become a fundamental component of deep learning models, including the field of computer vision, where spatial understanding plays a crucial role. A wide range of attention mechanisms have been designed to enable models to focus on the most relevant parts of input images, leading to improved performance and enhanced interpretability of the models' decision-making processes.[4–7]

The origins of attention in deep learning can be traced back to the revolutionary works in natural language processing (NLP), introduced as a way to address the limitations of traditional sequence-to-sequence models,[1] such as the inability to effectively capture long-range dependencies. The initial success of attention mechanisms in NLP tasks, such as machine translation and language modeling,[8–10] inspired researchers to explore their potential in other domains, including computer vision.

In the computer vision domain, attention mechanisms have been adapted to various tasks, such as image classification, object detection, and semantic segmentation. The key idea behind these attention mechanisms is to allow the model to dynamically focus on the most relevant spatial regions of the input image, rather than treating all regions equally. The selective focus has been shown to improve the model's performance and provide better interpretability of its decision-making.[11]

London Journal of Research in Computer Science & Technology

However, the traditional approaches to attention mechanisms in computer vision often suffer from a lack of smoothness in the attention maps, resulting in sharp transitions that negatively affect model generalization. This problem of spatial incoherence is particularly pronounced in the task of semantic segmentation, where accurate pixel-level predictions require a detailed understanding of the spatial relationships within the image.[12]

Semantic segmentation, the task of assigning a semantic label to each pixel in an image, relies heavily on capturing fine-grained details and understanding the spatial context. The existing attention mechanisms struggle to effectively capture the detailed relationships within the image due to the abrupt pixel-level transitions in the attention maps.[13, 14] This inconsistency in the attention maps often leads to inaccurate segmentation boundary predictions and a tendency for the model to overfit on the training data, limiting its ability to generalize well on unseen images.

Additionally, the lack of smoothness in attention maps can make models susceptible to noise.[15] Small perturbations in the input image may drastically alter the attention weights, causing unstable predictions.[16, 17] Such sensitivity to noise is particularly problematic in real-world scenarios where images are often corrupted by artifacts or imperfections.

To address these limitations, we present a new approach called Smooth Attention that incorporates a smoothness constraint, encouraging gradual changes in attention weights and mitigating the risks of sharp transitions or noise sensitivity. By enabling a detailed understanding of spatial relationships within the image, Smooth Attention leads to higher-level performance on tasks where spatial coherence is crucial, such as image semantic segmentation.

1.1 Related Work

The Smooth Attention methodology builds on the successful implementation of attention mechanisms in computer vision, particularly within transformer-based architectures such as the Vision Transformer[18] and Swin Transformer.[19] Our primary focus is the enhancement of the spatial coherence within a single attention layer, drawing inspiration from attention integration techniques in convolutional neural networks (CNNs), including Non-local Neural Networks[20] and the Convolutional Block Attention Module (CBAM).[21]

Spatial coherence has been a focal point in various applications, including image segmentation with CRF-RNN[22] and image generation through PatchGAN.[23] Smooth Attention introduces adaptive computation elements akin to Adaptive Computation Time[24] for recurrent neural networks (RNNs), selectively applying smoothing techniques.

Our framework is inspired by recent innovations in Attention Augmented Convolutional Net- works[25] and Focal Self-attention.[26] By leveraging the Chebyshev distance, we align with the principle of dynamically adjusting kernel weights based on local features during convolutional operations.[27]

The spatial coherence component of our approach resonates with the foundational concepts of CoordConv,[28] as it explicitly encodes spatial information into convolutional layers. Moreover, our methodology exhibits parallels with the Performer architecture,[29] which reduces computational complexity while maintaining high performance by approximating attention through random feature maps.

Building on the principles of adaptive filtering as outlined in[30] by Solomon et al., we effectively combine original and smoothed attention scores based on identified variations. Such selective smoothing enhances the established concept of gating in GRCNN,[31] enabling precise control over the influence of neighboring attention scores and pixels while preserving critical features.

II. METHODOLOGY

In this section, we present a new attention mechanism in convolutional neural networks called Smooth Attention. The module is designed to enhance the spatial coherence of attention maps while maintaining the flexibility and power of traditional attention mechanisms.[6, 32, 33] Our approach addresses the issue of inconsistent or "noisy" attention patterns that oftentimes arise in standard attention mechanisms, particularly in vision tasks where spatial coherence is crucial, like image segmentation of complex-shaped objects.[8, 34]

2.2 Architecture Overview

The Smooth Attention approach is implemented as a neural network module. It consists of three main components:

1. Query, key, and value projections (3.1.1)
2. Attention computation (3.1.2)
3. Smoothness enforcement mechanism (3.1.3)

Each of these components plays a crucial role in achieving the final goal of spatially coherent attention.

3.3 Query, Key, and Value Projections

The first step in the Smooth Attention module is the projection of the input tensor X into query (Q), key (K), and value (V) spaces. The projection is achieved using 1x1 convolutions (1), (2), (3).

$$Q = Conv1x1(X) \tag{1}$$

$$K = Conv1x1(X) \tag{2}$$

$$V = Conv1x1(X) \tag{3}$$

Where [Conv1x1] represents a convolutional layer with a 1x1 kernel. The final projections serve multiple purposes:

1. Allow the network to learn representations of the input that are suitable for computing attention.
2. Enable the module to adjust the channel dimensionality, reducing computational complexity.
3. Provide a learned transformation that emphasizes or suppresses aspects of the input for attention computation.

2.4 Attention Computation

After obtaining the projected tensors, we compute the attention scores using the scaled dot-product attention mechanism, similar to the one used in Transformer architectures,[1] shown in (4):

$$\text{attention} = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \tag{4}$$

Where:

4. QK^T represents the matrix multiplication of Q and the transpose of K;
5. d_k is the dimensionality of the key vectors;
6. $\sqrt{d_k}$ is used for scaling to counteract the effect of large dot products in high dimensions;
7. Softmax is applied to normalize the attention scores.

The computation results in the attention map where each position attends to all other positions, capturing global dependencies in the input.

2.1.3 Smoothness Enforcement Mechanism

The key concept in the Smooth Attention module is the introduction of a smoothness enforcement mechanism. The mechanism is designed to address the spatial incoherence in vision attention when nearby pixels may have radically different attention weights.

To enforce the smoothness, we compute the Chebyshev distance between each pixel's attention distribution and those of its eight neighboring pixels. The Chebyshev distance is defined as the maximum absolute difference across all dimensions (5).

$$d\_\text{chebyshev}(p, q) = max_i |p_i - q_i| \tag{5}$$

We compute this distance for each of the eight neighbors and take the maximum:

$$\text{max\_distance} = \max\left(d\_\text{chebyshev}(\text{attention}[i, j], \text{attention}[i + d_i, j + d_j])\right) \tag{6}$$

for d_i, d_j in $[(-1, -1), (-1, 0), (-1, 1), (0, -1), (0, 1), (1, -1), (1, 0), (1, 1)]$.

The choice of Chebyshev distance over other metrics is motivated by its sensitivity to the largest difference in any single dimension, which aligns well with the goal of detecting abrupt changes in attention patterns.[35, 36]

2.2 Smoothness Thresholding and Mask Creation

If the maximum Chebyshev distance exceeds a predefined threshold (a value between 0 and 1), we consider the attention at that pixel to be non-smooth. We create a binary mask where 1 indicates non-smooth regions and 0 indicates smooth regions (7).

$$\text{smoothing mask} = (\text{max chebyshev distance} > \text{threshold}).\text{float}() \tag{7}$$

2.3 Attention Application with Smoothness Consideration

The smoothing mask is then used to selectively apply the attention mechanism. In smooth regions, we allow the full attention mechanism to operate, while in non-smooth regions, we reduce the influence of the attention mechanism:

$$\text{output} = \gamma * \text{attention} * V + (1 - \gamma * \text{smoothing\_mask}) * x \tag{8}$$

Where:

- γ is a learnable parameter initialized to zero;
- V is the value projection of the input;
- x is the original input.

The learnable parameter γ allows the network to gradually incorporate the attention mechanism as training progresses. A soft start is used to help stabilize the training process and allow the network to learn "when and how much" to rely on the attention mechanism.

Due to its structure, Smooth Attention module offers several advantages:

1. Spatial Coherence: By enforcing local smoothness, the module encourages the production of more coherent attention maps, improving model's performance on complex computer vision tasks.
2. Noise Reduction: The smoothness constraint acts as a form of regularization, reducing the impact of noise and spurious correlations in the attention computation.
3. Interpretability: Smoother attention maps are more interpretable, providing clearer insights into which parts of the input the model is focusing on.
4. Adaptive Mechanism: The learnable parameter γ and the smoothness threshold provide additional level of engineering flexibility, allowing the network to adapt the strength of the smoothness constraint based on the task and data.

The Smooth Attention module extends traditional attention mechanisms by incorporating a local smoothness constraint. The constraint is enforced through applying the Chebyshev distances in the attention space, resulting in more coherent and accurate attention maps for convolutional neural networks. The module's design allows for a flexible trade-off between global attention capabilities and local coherence, making it adaptable to a wide range of vision tasks dependent on the spatial coherence level.

III. EXPERIMENTS

To demonstrate the effectiveness of Smooth Attention, we perform experiments on five diversified image segmentation datasets, including the Caltech-UCSD Birds-200-2011,[37] Large-Scale Dataset for Segmentation and Classification,[38] Fire Segmentation Image Dataset,[39] Kvasir-Instrument: Diagnostic and Therapeutic Tool Segmentation Dataset in Gastrointestinal Endoscopy,[40] and Flood Semantic Segmentation Dataset.[41]

To explore the influence of the attention module, we implement a custom model with the U-Net architecture[42–44] which is a popular choice for semantic segmentation tasks. We use ResNet18[45] as the encoder, which is followed by the Smooth Attention mechanism to help the model focus on important features. The decoder is implemented as a series of transposed convolutions that upscale the feature map.[46] Each transposed convolution is followed by a ReLU activation, except for the last one. The decoder gradually increases the spatial dimensions while reducing the number of channels.[47, 48] The final layer outputs the same number of channels as the number of classes for segmentation. The complete model architecture is demonstrated in Figure 1.

We perform experiments across multiple smoothness thresholds (from 0.1 to 0.9 with a 0.1 step) for each dataset, and compare the results with a threshold of 2.0 that illustrates the model without attention mechanism in place (any value above 1.0 means no smoothing constraint applied). All datasets are split into 80% for training, and 20% for testing. We choose IoU, Dice Coefficient, Test Accuracy, Test Precision Test F1 Score as metrics to compare the final results.[49]

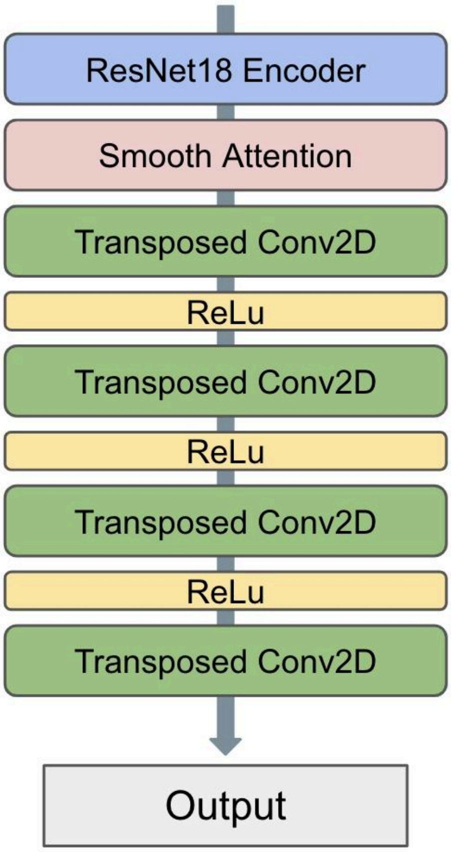


Fig. 1: Model Architecture.

Table 1: Experiments for Caltech-UCSD Birds-200-2011 results.

Metric	Test IOU	Test Dice	Test Acc	Test Prec	Test Recall	Test F1
0.1	0.7933	0.9231	0.9686	0.7674	**0.9838**	0.8550
0.2	**0.8028**	**0.9294**	0.9670	0.7457	0.9829	0.8430
0.3	0.7997	0.9292	0.9672	0.7641	0.9829	0.8521
0.4	0.7894	0.9186	0.9679	0.7654	0.9799	0.8495
0.5	0.7963	0.9258	0.9672	0.7609	0.9827	0.8494
0.6	0.7949	0.9260	0.9693	0.7740	0.9812	0.8562
0.7	0.7952	0.9239	0.9677	0.7685	0.9811	0.8526
0.8	0.8003	0.9272	0.9678	0.7700	0.9808	0.8493
0.9	0.7994	0.9229	**0.9699**	**0.7881**	0.9823	**0.8610**
2.0	0.7909	0.9222	0.9667	0.7497	0.9827	0.8453

The results for Caltech-UCSD Birds-200-2011[37] can be observed in Table 1. The superior performance of the Smooth Attention mechanism at lower smoothness thresholds (0.1, 0.2, 0.3) compared to the model without attention (threshold of 2.0) suggests that the attention module is effectively helping the model in focusing on the most relevant features in the image segmentation task. By applying the smoothness constraint, the attention mechanism is able to selectively high- light most informative regions of the input, leading to better segmentation accuracy, IoU, and Dice coefficient metrics.

We see the consistently high Test Recall values across all smoothness thresholds, showing that the model is able to correctly identify most of the positive instances in the test set, regardless of the attention mechanism's configuration. Similar to IoU, the Dice coefficient peaks at a lower smoothness threshold, demonstrating that low levels of attention smoothness provide better balance between focus and flexibility in feature selection. At the same time, the highest accuracy and pre- cision occur at a higher threshold, suggesting that for overall pixel-wise classification a smoother attention map reduces the noise and the number of false positives. The recall peak occurs at the lowest smoothness threshold, since true positives can be captured at a strict threshold easier. By applying a stricter smoothness constraint, the attention module is able to capture important visual cues, maximizing the model's ability to detect positive instances in the data.

The experiment results for the Large-Scale Dataset for Segmentation and Classification[38] can be observed in Table 2. The stronger IoU and Dice coefficient performance at lower smoothing thresholds indicates that the model is more effective in identifying variances in the input image data at strict thresholds. Such behavior unveils the rich tapestry of diversity within the fish seg- mentation data, highlighting the necessity for fine-grained approaches to capture the complexities of the images.

Table 2: Large-Scale Dataset for Segmentation and Classification.

Metric	Test IOU	Test Dice	Test Acc	Test Prec	Test Recall	Test F1
0.1	0.9344	0.9764	0.9882	0.9426	0.9904	0.9618
0.2	**0.9364**	0.9784	**0.9887**	0.9434	**0.9916**	**0.9638**
0.3	0.9309	0.9749	0.9879	0.9493	0.9901	0.9600
0.4	0.9320	0.9751	0.9883	0.9399	0.9898	0.9627
0.5	0.9355	**0.9787**	0.9886	0.9426	0.9910	0.9630
0.6	0.9283	0.9706	0.9874	0.9383	0.9872	0.9602
0.7	0.9247	0.9694	0.9886	**0.9519**	0.9866	0.9620
0.8	0.9290	0.9731	0.9885	0.9434	0.9882	0.9626
0.9	0.9336	0.9759	0.9879	0.9465	0.9899	0.9612
2.0	0.9363	0.9781	0.9886	0.9436	0.9915	0.9611

Consistently high accuracy scores for all thresholds, including 2.0 (no attention), showcase the absence of background complexity within image data, helping the model be accurate in object segmentation with and without smoothing. Still, slight improvements can be seen when attention smoothing is applied with peak scores seen at thresholds of 0.2, 0.5 and 0.7.

Precision metric demonstrates the best result at a threshold of 0.7, showing accurate prediction of pixel-level distribution. The higher thresholds can be more beneficial for tasks requiring broader feature recognition helped by greater noise suppression and lesser sensitivity to variation in input data. Moreover, high values of Recall and F1 Score at lower thresholds are a clear representation of the model's effectiveness in capturing critical details due to strict smoothing masks.

Table 3 demonstrates the results for the Fire Segmentation Image Dataset.[39] The IoU and Dice coefficient peaked at a smoothing threshold of 0.4, with strong performance also observed at 0.5 and 0.6. This behavior can be attributed to the special characteristics of fire images in the dataset.

Table 3: Fire Segmentation Image Dataset

Metric	Test IOU	Test Dice	Test Acc	Test Prec	Test Recall	Test F1
0.1	0.8484	0.9228	0.9904	0.4903	0.9630	0.6322
0.2	0.8512	0.9246	0.9913	0.4868	0.9633	0.6403
0.3	0.8488	0.9189	**0.9925**	**0.5269**	0.9599	0.6377
0.4	**0.8518**	**0.9277**	0.9914	0.4907	0.9662	0.6446
0.5	0.8501	0.9248	0.9910	0.4753	0.9648	0.6329
0.6	0.8507	0.9268	0.9919	0.5096	**0.9656**	0.6489
0.7	0.8498	0.9231	0.9918	0.5027	0.9628	0.6460
0.8	0.8480	0.9208	0.9918	0.5057	0.9594	**0.6520**
0.9	0.8497	0.9231	0.9915	0.4888	0.9608	0.6314
2.0	0.8504	0.9211	0.9911	0.4837	0.9623	0.6399

Fire scenes typically exhibit complex, irregular shapes with varying intensity and color gradients, making them challenging to segment accurately.

The dataset's nature, featuring dynamic fire boundaries and potential smoke interference, explains why moderate smoothing thresholds (0.4 - 0.6) outperform others. At these levels, the algorithm effectively balances detail preservation and noise reduction. Lower thresholds (0.1 - 0.3) likely retain too much noise and small, irrelevant features, while higher thresholds (0.7 - 0.9) risk oversimplifying the fire's complex structure.

Consistently high values of Accuracy and Recall suggest the model's ability to correctly identify true positives and relevant instances in the image dataset, regardless of whether smoothing is applied. However, attention smoothing does lead to marginally improved performances in these metrics.

The balanced approach to smoothing not only improves traditional segmentation metrics like IoU and Dice coefficient but also enhances the model's overall predictive capabilities across a broader range of performance indicators.

Table 4: Kvasir-Instrument: Diagnostic and Therapeutic Tool Segmentation Dataset in Gastrointestinal Endoscopy

Metric	Test IOU	Test Dice	Test Acc	Test Prec	Test Recall	Test F1
0.1	0.9105	0.9405	0.9853	0.9427	0.9370	0.9138
0.2	**0.9248**	**0.9536**	**0.9881**	0.9489	**0.9674**	**0.9294**
0.3	0.9223	0.9504	0.9876	**0.9632**	0.9548	0.9263
0.4	0.9071	0.9348	0.9852	0.9231	0.9367	0.9136
0.5	0.9125	0.9420	0.9848	0.9162	0.9364	0.9163
0.6	0.9054	0.9371	0.9847	0.9077	0.9326	0.9167

0.7	0.9052	0.9374	0.9828	0.9121	0.9345	0.9081
0.8	0.9093	0.9394	0.9861	0.9083	0.9587	0.9217
0.9	0.9091	0.9390	0.9857	0.9115	0.9470	0.9197
2.0	0.9108	0.9376	0.9855	0.9152	0.9479	0.9208

Table 4 presents the results of the Kvasir-Instrument: Diagnostic and Therapeutic Tool Seg- mentation Dataset in Gastrointestinal Endoscopy.[40] The dataset contains images of the commonly used gastrointestinal endoscopy surgical tools that pose challenges for segmentation models due to the shape complexity of the surgical tools captured in the scenes.

The best performance across numerous evaluation metric criteria can be seen at lower thresholds (0.1-0.3) with peak values achieved at a threshold of 0.2. Such behavior clearly indicates the importance of fine-grained attention in the medical industry where the decision based on spatial understanding of the structure of organ imagery and surgical tools can play a major role in life and death situations.

In tasks of segmentation and classification the medical instruments are commonly segmented incorrectly and mistaken for others. The lower attention threshold strictly punishes false positives, leading to more accurate segmentation results, as seen at the metrics-performance level.

Additionally, the high values of Dice and IOU at the threshold of 0.2 show that the model is able to capture image-specific details and identify accurately both linear and non-linear segmentation boundaries. At the same time, high values of Accuracy, Precision, Recall and F1 Score at the lower thresholds highlight the model's readiness to adhere to local sensitivity in pixel-level variations that are crucial in the medical field.

Table 5: Flood Semantic Segmentation Dataset

Metric	Test IOU	Test Dice	Test Acc	Test Prec	Test Recall	Test F1
0.1	0.8746	0.9321	0.9571	0.9450	0.9669	**0.9515**
0.2	0.8694	0.9274	0.9542	0.9521	**0.9790**	0.9486
0.3	0.8765	0.9309	0.9543	0.9432	0.9712	0.9480
0.4	**0.8772**	**0.9334**	0.9554	**0.9595**	0.9662	0.9501
0.5	0.8374	0.9309	0.9560	0.9556	0.9691	0.9495
0.6	0.8728	0.9294	0.9531	0.9501	0.9658	0.9470
0.7	0.8744	0.9317	0.9555	0.9452	0.9612	0.9502
0.8	0.8674	0.9263	0.9522	0.9419	0.9622	0.9446
0.9	0.8696	0.9292	0.9518	0.9517	0.9646	0.9469
2.0	0.8720	0.9308	0.9542	0.9514	0.9605	0.9468

Table 5 reveals the results for the Flood Semantic Segmentation Dataset.[41] The dataset covers a wide range of image data of flood disaster area sensing in multiple geographical regions, introducing the complexity of segmentation variation.

The optimal performance can be observed at the thresholds (0.4-0.6) with strong IOU, Dice and Precision results. Due to the fluid image structure variations, it is important to find the balancing between border smoothing and fine-grained detail retention. Such balance ensures the model cap- tures segmentation features while minimizing artifacts arising from overly aggressive smoothing techniques.[49, 50]

Lower thresholds yield high recall but lower precision values. This phenomenon occurs because the balance achieved ensures the detection of image details, resulting in a higher percentage of true positives; however, still providing the needed leniency towards false positives, as the complexity of flooded areas in varying geographical terrains—such as reflections, shallow water, and debris—can cause the model to overreact to minor pixel distributions variations.

Higher thresholds tend to perform less effectively due to a tendency to under-segment the shape complexities of flooded and non-flooded areas. This loss of detail obscures critical features, making it challenging to accurately distinguish between waterlogged regions and surrounding terrain. Consequently, important contextual information is lost, leading to inaccuracies in assessing the extent of flooding.

We demonstrate the comparison of the attention values as heatmaps for the Flood Semantic Segmentation Dataset [41] at different thresholds in Figures 2 and 3. Figure 2 illustrates attention heatmap at a threshold of 2.0 (without smoothing applied), resulting in a more fragmented representation of attention values across the map. This higher threshold leads to sharper and isolated areas of focus, which obscure subtle relationships in the image data. Conversely, Figure 3 presents attention heatmap at a threshold of 0.4 (with effective smoothing applied), showing a more cohe- sive and visually integrated attention values distribution. The smoothing process creates a gradual transition between areas of high and low attention, enhancing spatial interpretability and resulting in a more stable segmentation model performance.

Developing further the insights from the previous figures, we introduce in Figures 4 and 5 the 3D heatmap representation of attention for the Flood Semantic Segmentation Dataset [41] at the same respective thresholds. Figure 4 shows the 3D heatmap at a threshold of 2.0 (without smoothing applied), clearly revealing a sparse and jagged attention landscape. The isolated peaks in this visualization indicate areas of high attention, but the overall structure appears disjointed, making it challenging to discern the concise relationships between different attention regions. In contrast, Figure 5 presents the 3D heatmap with a threshold of 0.4 (effective smoothing applied), which significantly alters the visual interpretation of attention. The smoothing process results in a more fluid attention surface, building a precise understanding of what regions help the model complete the task accurately. After the smoothing constraint is applied, the attention regions shift due to the averaging effect that reduces noise and emphasizes broader trends in the values' distribution. Such cohesive representation not only enhances the visibility of region-level importance but also facilitates the identification of underlying patterns within the image data.[51]

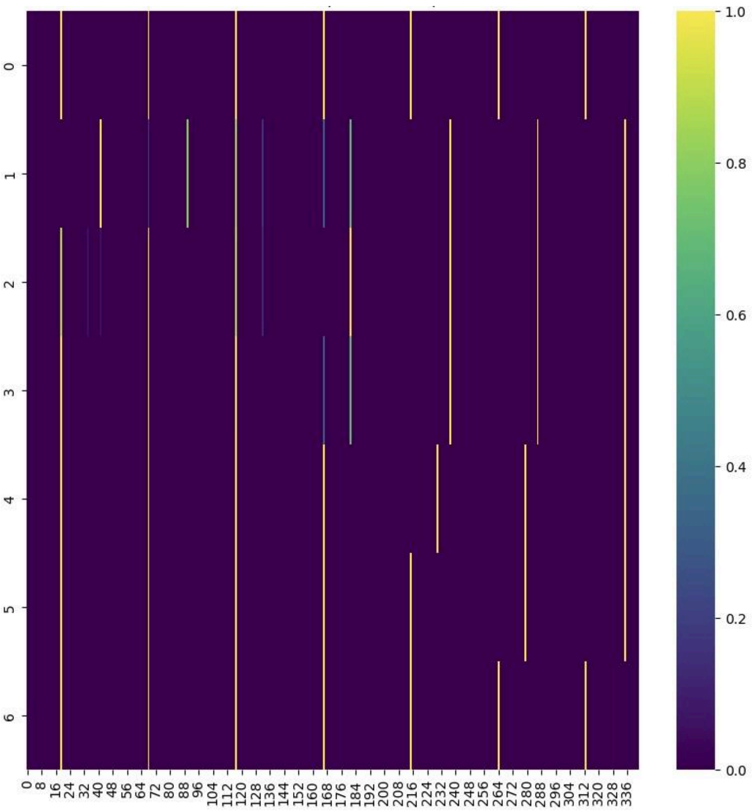


Fig 2: Attention as a Heatmap for Flood Semantic Segmentation Dataset at threshold 2.0 (without smoothing).

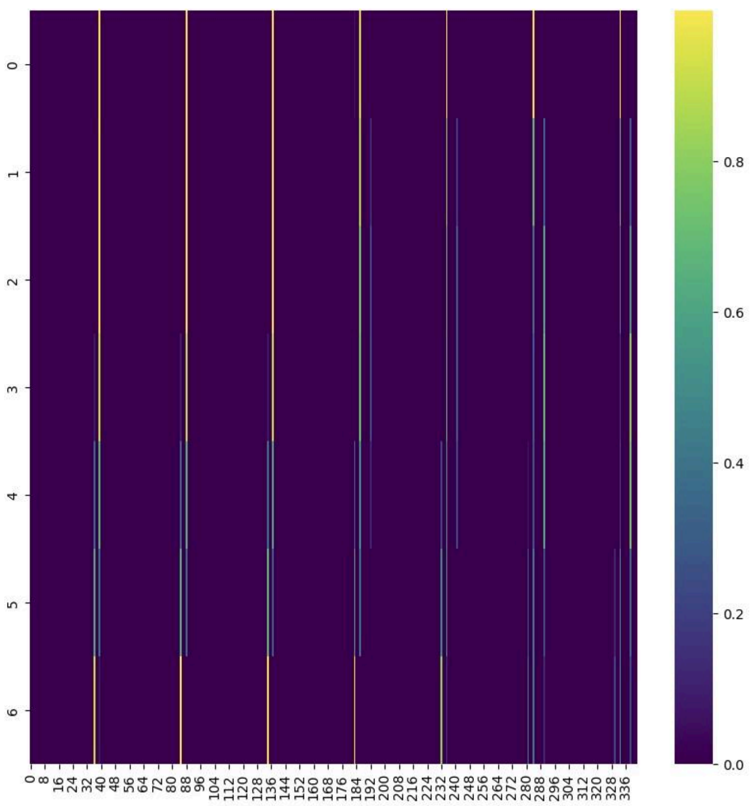


Fig 3: Attention as a Heatmap for Flood Semantic Segmentation Dataset at threshold 0.4 (with smoothing).

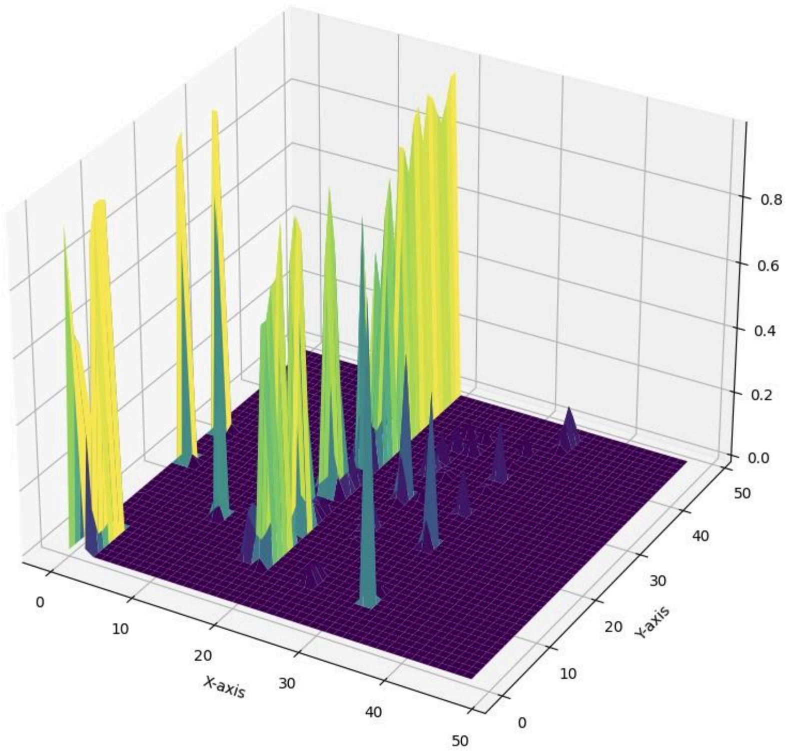


Fig. 4: Attention as a 3D Heatmap for Flood Semantic Segmentation Dataset at threshold 2.0 (without smoothing).

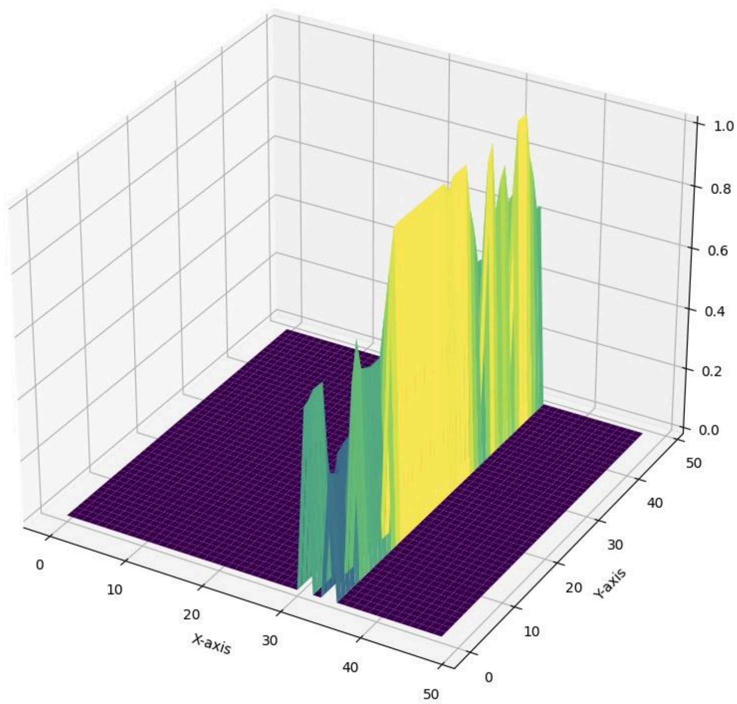


Fig. 5: Attention as a 3D Heatmap for Flood Semantic Segmentation Dataset at threshold 0.4 (with smoothing).

Figures 6 and 7 explore the representation of attention values using 3D scatter plots, providing a distinct perspective on the attention distribution under varying levels of smoothing for the Large- Scale Dataset for Segmentation and Classification [38]. Figure 6 illustrates the attention values at a threshold of 0.9 (with light smoothing applied). It can be observed that the attention points are relatively sparse, with some clusters indicating areas of significant focus. However, higher threshold values limit the mechanism's ability to identify finer details in the distribution map. The light smoothing enhances the overall visual coherence, developing a stronger understanding of the focusing area where the fish is illustrated, yet it still retains some of the original fragmentation.

In contrast, Figure 7 presents the attention values at a threshold of 0.1 (with strong smoothing applied). Such an approach results in a dense interconnected scatter plot, where attention points are uniformly distributed across the attention region. The strong smoothing effectively blurs the boundaries between areas of high and low attention, creating a continuous representation of attention across the image data. Figure 7 highlights the strict relationships between different regions, making it easier to identify patterns that could have been missed in the original attention map's spatial inconsistency.

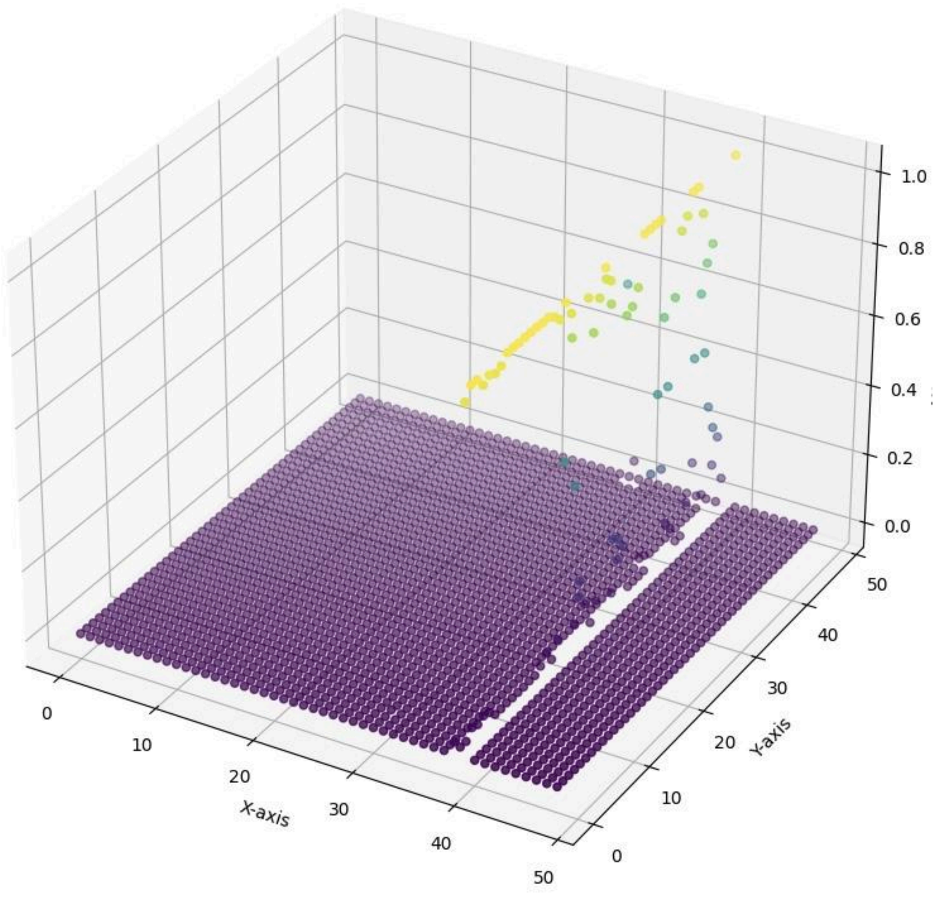


Fig. 6: 3D Scatter Plot for Attention Values at threshold 0.9 (light smoothing) for Large-Scale Dataset for Segmentation and Classification.

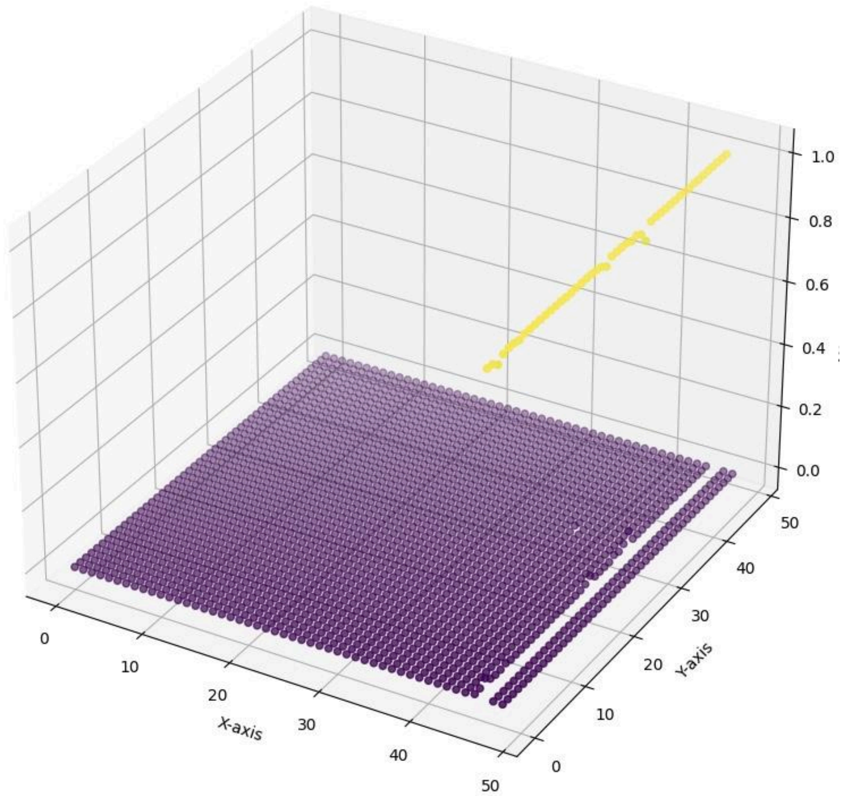


Fig. 7: 3D Scatter Plot for Attention Values at threshold 0.1 (strong smoothing) for Large-Scale Dataset for Segmenta- tion and Classification.

Figures 8 and 9 present the final predicted segmentation results for the Flood Semantic Seg- mentation Dataset [41]. Figure 8 illustrates the segmentation output at a threshold of 2.0 (withoutsmoothing applied), revealing the segmented images that are characterized by abrupt edges and fragmented regions. While some areas of interest are accurately captured, the lack of smooth- ing leads to a disjointed segmentation that does not effectively represent the underlying structural details in the spatial data.

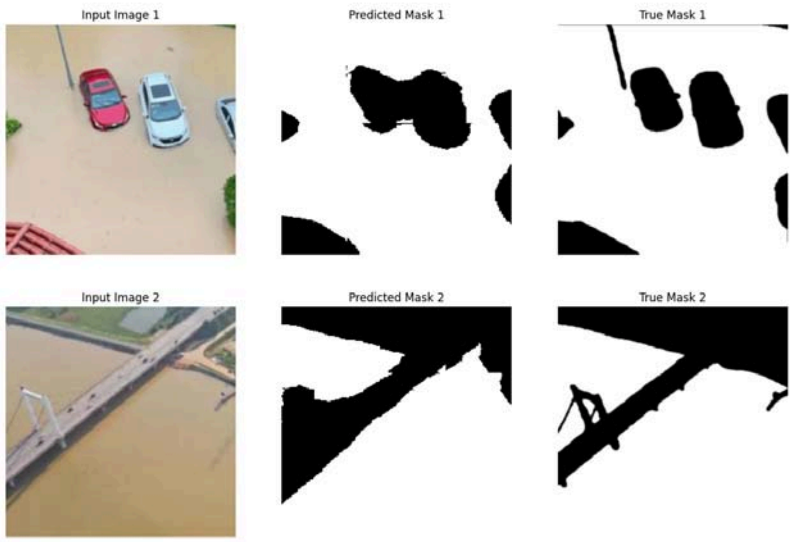


Fig. 8: Predicted Segmentation Result threshold 2.0 (without smoothing) for Flood Semantic Segmentation Dataset.

At the same time Figure 9 demonstrates the predicted segmentation mask at a threshold of 0.4 (with effective smoothing applied). The result with the Smooth Attention method achieves a more coherent and unified segmentation, having smoother transitions between the object boundaries. The application of smoothing constraint enhances the model's ability to capture complex shapes and relationships on the pixel-level, resulting in a segmentation mask that outlines detailed variation within the images.

We can see that the introduction of Smooth Attention notably improves the spatial distribution of the attention values within the attention map, helping the model to achieve better segmentation results. By incorporating a smoothness constraint, the Smooth Attention method encourages gradual changes in attention weights, which effectively mitigates the noise sensitivity problems with attention value distribution. We achieve the reduction of sharp transitions and foster a deeper understanding of spatial relationships of the image data.

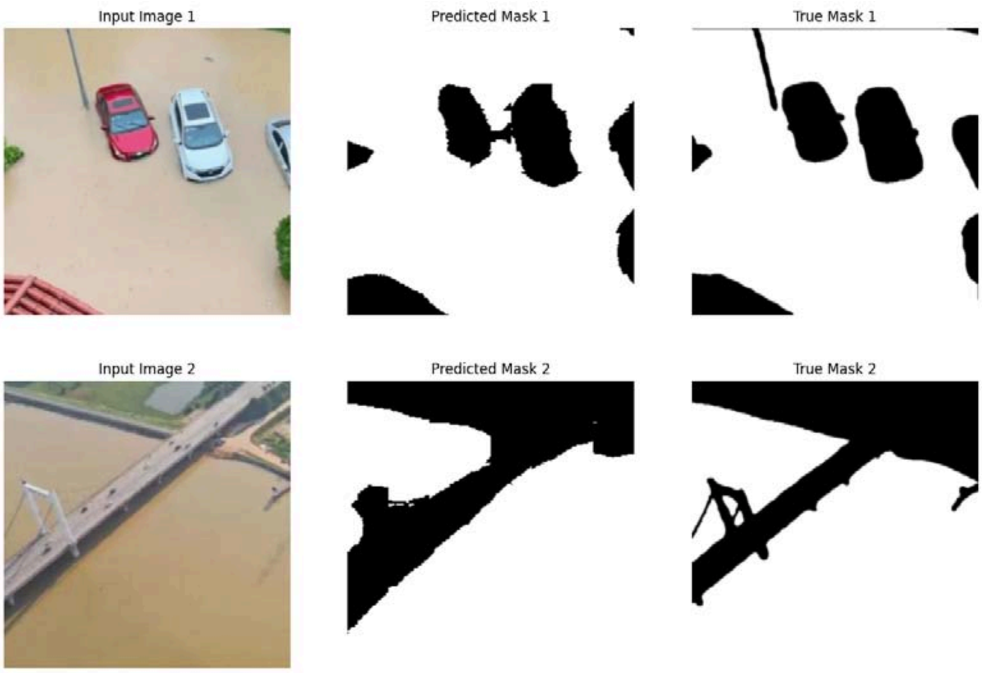


Fig. 9: Predicted Segmentation Result threshold 0.4 (with smoothing) for Flood Semantic Segmentation Dataset.

IV. CONCLUSION

To summarize, we introduce a new attention approach called Smooth Attention, designed to en- hance the spatial coherence of attention maps in convolutional neural networks, particularly for vision tasks like image segmentation. Our experiments demonstrate that by incorporating a spatial distribution-aware smoothness enforcement mechanism, we improve the quality of the model's focus on relevant regions of the input images.

The Smooth Attention module effectively mitigates the complications of noisy attention patterns, resulting in smoother attention maps that still maintain the detail awareness inherent in traditional attention mechanism ideas. By leveraging Chebyshev distance to enforce spatial-aware smoothness, we achieve a balance between global and local attention, enhancing both the model's performance and interpretability.

Moreover, the adaptive nature of the learnable parameter and the tunable smoothness threshold provide the flexibility needed to tailor the mechanism to specific tasks and datasets. Such adaptability is essential for applications where varying degrees of spatial coherence are required, allowing to optimize the performance for the particular use cases based on the metrics results.

Future work will explore the application of Smooth Attention across other datasets and tasks, extending its use beyond image segmentation to other domains such as object detection and image captioning.

We believe that the Smooth Attention mechanism holds promise for advancing the interpretabil- ity and effectiveness of attention in convolutional neural networks for computer vision tasks. By addressing the challenge of spatial incoherence in attention maps, our approach paves the way for models that can better understand the complex visual data. We anticipate that this work will inspire further research in attention mechanisms in computer vision that prioritize interpretability and coherence, ultimately leading to more trustworthy AI systems in critical applications.

Disclosures

All the authors admit no conflict of interests.

Code, Data, and Materials Availability

All the datasets used are publicly cited and available. The code is submitted to GitHub storage and can be accessed publicly at: https://github.com/sparcus-technologies/Smooth- Attention-Improving-Image-Semantic-Segmentation

ACKNOWLEDGMENTS

The authors express gratitude to Sparcus Technologies Limited for providing computational re- sources for this research.

REFERENCES

1. Vaswani, A. (2017). Attention is all you need. Advances in Neural Information Processing Systems.
2. Galassi, A., Lippi, M., & Torroni, P. (2020). Attention in natural language processing. IEEE transactions on neural networks and learning systems, 32(10), 4291-4308.
3. Guo, M. H., Xu, T. X., Liu, J. J., Liu, Z. N., Jiang, P. T., Mu, T. J., ... & Hu, S. M. (2022). Attention mechanisms in computer vision: A survey. Computational visual media, 8(3), 331-368.
4. Yang, X. (2020, December). An overview of the attention mechanisms in computer vision. In Journal of physics: Conference series (Vol. 1693, No. 1, p. 012173). IOP Publishing.
5. Wang, F., & Tax, D. M. (2016). Survey on the attention based RNN model and its applica- tions in computer vision. arXiv preprint arXiv:1601.06823.
6. Bello, I., Zoph, B., Vaswani, A., Shlens, J., & Le, Q. V. (2019). Attention augmented convo- lutional networks. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 3286-3295).
7. Ramachandran, P., Parmar, N., Vaswani, A., Bello, I., Levskaya, A., & Shlens, J. (2019). Stand-alone self-attention in vision models. Advances in neural information processing sys- tems, 32.
8. Luong, M. T. (2015). Effective approaches to attention-based neural machine translation. arXiv preprint arXiv:1508.04025.
9. Zhang, B., Xiong, D., – Su, J. (2018). Neural machine translation with deep attention. IEEE transactions on pattern analysis and machine intelligence, 42(1), 154-163.
10. Maruf, S., Martins, A. F., & Haffari, G. (2019). Selective attention for context-aware neural machine translation. arXiv preprint arXiv:1903.08788.
11. You, Q., Jin, H., Wang, Z., Fang, C., & Luo, J. (2016). Image captioning with semantic attention. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4651-4659).

12. Sun, J., Jiang, J., & Liu, Y. (2020, December). An introductory survey on attention mecha- nisms in computer vision problems. In 2020 6th International Conference on Big Data and Information Analytics (BigDIA) (pp. 295-300). IEEE.
13. Li, H., Xiong, P., An, J., & Wang, L. (2018). Pyramid attention network for semantic segmentation. arXiv preprint arXiv:1805.10180.
14. Guo, M. H., Lu, C. Z., Hou, Q., Liu, Z., Cheng, M. M., & Hu, S. M. (2022). Segnext: Rethinking convolutional attention design for semantic segmentation. Advances in Neural Information Processing Systems, 35, 1140-1156.
15. Konate, S., Lebrat, L., Santa Cruz, R., Bourgeat, P., Doreˆ, V., Fripp, J., ... & Salvado, O. (2021, April). Smocam: Smooth conditional attention mask for 3d-regression models. In 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) (pp. 362-366). IEEE.
16. Yao, Y., Ren, J., Xie, X., Liu, W., Liu, Y. J., & Wang, J. (2019). Attention-aware multi- stroke style transfer. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 1467-1475).
17. Jiang, P. T., Han, L. H., Hou, Q., Cheng, M. M., & Wei, Y. (2021). Online attention accumulation for weakly supervised semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(10), 7062-7077.
18. Alexey, D. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv: 2010.11929.
19. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10012-10022).
20. Wang, X., Girshick, R., Gupta, A., & He, K. (2018). Non-local neural networks. In Proceed- ings of the IEEE conference on computer vision and pattern recognition (pp. 7794-7803).
21. Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). Cbam: Convolutional block attention module. In Proceedings of the European conference on computer vision (ECCV) (pp. 3-19).
22. Zheng, S., Jayasumana, S., Romera-Paredes, B., Vineet, V., Su, Z., Du, D., ... & Torr, P. H. (2015). Conditional random fields as recurrent neural networks. In Proceedings of the IEEE international conference on computer vision (pp. 1529-1537).
23. Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with condi- tional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1125-1134).
24. Graves, A. (2016). Adaptive computation time for recurrent neural networks. arXiv preprint arXiv:1603.08983.
25. Bello, I., Zoph, B., Vaswani, A., Shlens, J., & Le, Q. V. (2019). Attention augmented convo- lutional networks. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 3286-3295).
26. Yang, J., Li, C., Zhang, P., Dai, X., Xiao, B., Yuan, L., & Gao, J. (2021). Focal self-attention for local-global interactions in vision transformers. arXiv preprint arXiv:2107.00641.
27. Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. Pattern recognition letters, 119, 3-11.
28. Liu, R., Lehman, J., Molino, P., Petroski Such, F., Frank, E., Sergeev, A., & Yosinski, J. (2018). An intriguing failing of convolutional neural networks and the coordconv solution. Advances in neural information processing systems, 31.
29. Choromanski, K., Likhosherstov, V., Dohan, D., Song, X., Gane, A., Sarlos, T., ... & Weller, A. (2020). Rethinking attention with performers. arXiv preprint arXiv:2009.14794.
30. Solomon, J., Crane, K., Butscher, A., & Wojtan, C. (2014). A general framework for bilat- eral and mean shift filtering. arXiv preprint arXiv:1405.4734, 1(2), 3.
31. Wang, J., & Hu, X. (2021). Convolutional neural networks with gated recurrent connections. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(7), 3421-3435.

32. Volz, S., Bruhn, A., Valgaerts, L., & Zimmer, H. (2011, November). Modeling temporal coherence for optical flow. In 2011 International Conference on Computer Vision (pp. 1116- 1123). IEEE.
33. Tong, X., Xu, R., Liu, K., Zhao, L., Zhu, W., & Zhao, D. (2023). A Deep-Learning Ap- proach for Low-Spatial-Coherence Imaging in Computer-Generated Holography. Advanced Photonics Research, 4(1), 2200264.
34. Zabih, R., & Kolmogorov, V. (2004, June). Spatially coherent clustering using graph cuts. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. (Vol. 2, pp. II-II). IEEE.
35. Li, F., Lebanon, G., & Sminchisescu, C. (2012, June). Chebyshev approximations to the his- togram X 2 kernel. In 2012 IEEE Conference on Computer Vision and Pattern Recognition (pp. 2424-2431). IEEE.
36. Koenderink, J., & van Doom, A. (1998). Shape from Chebyshev nets. In Computer Vi- sion—ECCV'98: 5th European Conference on Computer Vision Freiburg, Germany, June 2–6, 1998 Proceedings, Volume II 5 (pp. 215-225). Springer Berlin Heidelberg.
37. Wah, C., Branson, S., Welinder, P., Perona, P., & Belongie, S. (2011). The caltech-ucsd birds-200-2011 dataset.
38. Ulucan, O., Karakaya, D., & Turkan, M. (2020, October). A large-scale dataset for fish segmentation and classification. In 2020 Innovations in Intelligent Systems and Applications Conference (ASYU) (pp. 1-5). IEEE.
39. DiversisAI. (n.d.). Fire segmentation image dataset. Kaggle. https://www.kaggle.com/ datasets/diversisai/fire-segmentation-image-dataset.
40. Jha, D., Ali, S., Emanuelsen, K., Hicks, S. A., Thambawita, V., Garcia-Ceja, E., Riegler, M. A., de Lange, T., Schmidt, P. T., Johansen, H. D., Johansen, D., & Halvorsen, P. (2021). Kvasir-Instrument: Diagnostic and therapeutic tool segmentation dataset in gastrointestinal endoscopy. In MultiMedia Modeling (pp. 218–229). Springer International Publishing.
41. Li, H. (n.d.). Flood semantic segmentation dataset. Kaggle. https://www.kaggle.com/ datasets/lihuayang111265/flood-semantic-segmentation-dataset.
42. Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V. (2021). U-net and its vari- ants for medical image segmentation: A review of theory and applications. IEEE access, 9, 82031-82057.
43. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomed- ical image segmentation. In Medical image computing and computer-assisted interven- tion–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer International Publishing.
44. Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., ... & Rueck- ert, D. (2018). Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.
45. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
46. Gao, H., Yuan, H., Wang, Z., & Ji, S. (2019). Pixel transposed convolutional networks. IEEE transactions on pattern analysis and machine intelligence, 42(5), 1218-1227.
47. Cho, K., Courville, A., & Bengio, Y. (2015). Describing multimedia content using attention- based encoder-decoder networks. IEEE Transactions on Multimedia, 17(11), 1875-1886.
48. Ji, Z., Xiong, K., Pang, Y., & Li, X. (2019). Video summarization with attention-based en- coder–decoder networks. IEEE Transactions on Circuits and Systems for Video Technology, 30(6), 1709-1717.
49. Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). A review on deep learning techniques applied to semantic segmentation. arXiv preprint arXiv:1704.06857.

50. Du, S., Fan, H., Zhao, M., Zong, H., Hu, J.,& Li, P. (2022). A two-stage method for single image de-raining based on attention smoothed dilated networks. IET Image Processing, 16(10), 2557-2567.
51. Ouyang, W., Zeng, X., & Wang, X. (2013). Modeling mutual visibility relationship in pedestrian detection. In Proceedings of the IEEE conference on computer vision and pat- tern recognition (pp. 3222-3229).

This page is intentionally left blank

The Extent to Which Graphic Designers in Oman Rely on AI Software

Laila Alsibani

Sultan Qaboos University

ABSTRACT

The current study examined the extent to which graphic designers in Oman rely on artificial intelligence software and it highlighted full-time and freelance designers specifically. Artificial intelligence has advanced tremendously in the last several years, and this has had an impact on human life in general. In other words, artificial intelligence is a tool that recognizes patterns and assists humans in problem-solving. From this perspective, AI can serve as a designer's aide. This study aims to measure the extent to which graphic designers use AI software to accomplish their design tasks and to identify the most frequently used AI programs among graphic designers in Oman. This research utilized a quantitative method, collecting data through an online survey questionnaire. By exploring the familiarity of designers with AI programs in graphic design within Oman, this study provides a foundational resource for future research. It is among the few studies that focus on this specific target group. Future research could investigate the long-term effects and evolving practices of users as AI tools become increasingly integrated into design workflows.

Keywords: artificial intelligence, design software, graphic design.

Classification: ACM Code: I.2.10

Language: English

Great Britain Journals Press

LJP Copyright ID: 975816
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science & Technology

Volume 24 | Issue 2 | Compilation 1.0

The Extent to Which Graphic Designers in Oman Rely on AI Software

Laila Alsibani

ABSTRACT

The current study examined the extent to which graphic designers in Oman rely on artificial intelligence software and it highlighted full-time and freelance designers specifically. Artificial intelligence has advanced tremendously in the last several years, and this has had an impact on human life in general. In other words, artificial intelligence is a tool that recognizes patterns and assists humans in problem-solving. From this perspective, AI can serve as a designer's aide. This study aims to measure the extent to which graphic designers use AI software to accomplish their design tasks and to identify the most frequently used AI programs among graphic designers in Oman. This research utilized a quantitative method, collecting data through an online survey questionnaire. By exploring the familiarity of designers with AI programs in graphic design within Oman, this study provides a foundational resource for future research. It is among the few studies that focus on this specific target group. Future research could investigate the long-term effects and evolving practices of users as AI tools become increasingly integrated into design workflows.

Keywords: artificial intelligence, design software, graphic design.

Author: Graphic Designer / Sultan Qaboos University.

I. INTRODUCTION

Artificial Intelligence (AI) involves creating advanced machines capable of performing specific tasks with greater proficiency than humans (Ali Elfa & Dawood, 2023). Artificial intelligence has advanced tremendously in the last several years, and this has had an impact on human life in general (al-Dulaimi, 2024). In other words, artificial intelligence is a tool that recognizes patterns and assists humans in problem-solving. From this perspective, AI can serve as a designer's aide (Huo & Wang, 2022). The use of computer-based software to support design modeling, analysis, review, and documentation is known as computer-aided design (CAD). The advantages of CAD can be significantly improved by integrating it with artificial intelligence (AI), augmented reality, and manufacturing technologies. AI can generate an intelligent graphics interface and transform laborious design processes into sophisticated ones (Hunde & Woldeyohannes,2022). Integrating artificial intelligence theory into computer-aided art design is an emerging research focus and a contemporary trend in the modernization of industrial design. This approach not only incorporates AI research findings into computer-aided art design but also broadens the scope of AI applications, creating a complementary relationship that fosters mutual advancement. The field of computer art has grown significantly with the advancement of artificial intelligence technology, and a vast amount of new works are produced every year (Deng & Chen, 2021).

Graphic design is not an exception to the fast-expanding field of artificial intelligence (AI), which is already starting to alter many other industries. The creation of new tools and methods that are revolutionizing the creation and experience of graphic design is a result of AI's capacity to learn from and comprehend enormous volumes of data (Mustafa, 2023). Design can be seen as an intentional act

London Journal of Research in Computer Science & Technology

of creation (Huo & Wang, 2022).In the modern world, graphic design has become the most significant industry since it improves user experiences across the board, including communication, marketing, product advertising, brand logo creation, game design, applications, and website layouts (Sindhura & Abdul,2021). By the way, Instead of being passive consumers of AI tools, graphic designers are active learners who are always improving their abilities to successfully integrate AI into their design processes (Munzier et al ., 2024). It's a field largely focused on inventiveness, imagination, and the capacity of the designer to produce the best designs (Sindhura & Abdul,2021).

II. PROBLEM STATEMENT

To keep pace with developments in technology and with the global trend in the use of artificial intelligence and its employment in various fields, this research focuses on one of the significant areas that are considered at the forefront in the use of artificial intelligence techniques and tools in Oman, which is graphic design. Artificial intelligence applications that serve the graphic designer are among the indispensable applications, especially since they play a major role in facilitating tasks and saving the designer's time. Therefore, it is necessary to shed light on the extent of the graphic designer's need and reliance on using artificial intelligence tools to complete his work. This study also reveals the most prominent artificial intelligence programs used by the designer to complete his design work. Thus, studying the designer's need to develop his skills in these particular programs based on his focus on them in professionally completing his tasks. In addition, that leads to enriching the design process and also offering unique and engaging visual content.

III. RESEARCH OBJECTIVES

The main goal of this research is to investigate the extent to which graphic designers in Oman need to resort to using artificial intelligence tools instead of non-AI tools software in the design process. In addition to highlighting the software that graphic designers prefer to use the most. Through the aforementioned objectives, this study addressed two questions. First, to what extent do graphic designers rely on AI software to accomplish their design enterprises? Second, what are the most used AI software by graphic designers in Oman?

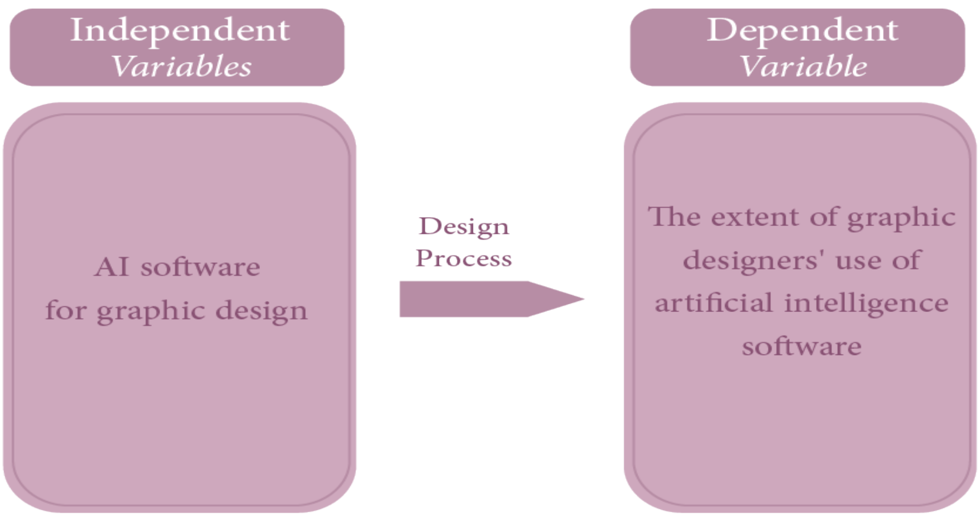


Fig. 1: Framework of the study

IV. LITERATURE REVIEW

4.1 Artificial Intelligence as a Design Material

Implementing AI in graphic design is a complicated process that involves major investment and careful consideration of ethical and technological constraints (Mustafa, 2023). Design is a reflective discipline in which designers engage in a "conversation with materials" to imagine things that do not currently exist. Materials "talk back" to designers, suggesting potential and limitations [...] previous research on AI as a design material revealed that designers who had created a high number of "designedly abstractions" of AI's capabilities were more successful and comfortable working with AI. These designers were able to engage in reflective conversations with AI, and they used data scientists' technical expertise to "talk back" on what is conceivable (Yildirim et al., 2022). AI technology is revolutionizing the creation of visual content by enhancing the accessibility and efficiency of design tools. AI-powered tools, such as AI Image Generators, allow for easy creation of high-quality graphics for marketing, entertainment, and personal usage. These programs use powerful algorithms to generate images based on user input, themes, styles, and preferences (Rivas, 2024).

4.2 AI Software in Graphic Design

While amateur graphic design tools primarily rely on templates, AI in professional graphic design software has tended to focus on automating arduous chores, which creatively driven professional graphic designers appreciate. For instance, Photoshop has long included "actions" to automatically apply duplicate editing effects to many photos in addition to features like the "healing brush" that make isolated, difficult compositing work easier. Professional graphic design software tools are still mostly unaffected by artificial intelligence (AI), except certain focused applications of the technology. Professional graphic design software does, in fact, often highlight its capacity to provide designers exact control over the creative process, including step-by-step adjustments to color management and type kerning. It is therefore clear that automating the more tiresome and repetitive "art working" duties associated with creating graphic designs would be beneficial to experts and provide them "more opportunity and time to concentrate on the creative side of projects " (Meron, 2022). By the way, The year 2023 has seen an unparalleled increase in the demand for AI design tools. Google Trends indicates that the amount of searches for software and tools linked to AI design has increased by 1700% just between 2022 and 2023 (Loewy, 2024).

The research community consists of a diverse group of artificial intelligence applications that are used in graphic design, and because of the large number of applications, the researcher intentionally chose five of them due to their diversity in graphic design work and because of their frequent use in graphic design alone. The design software chosen in this research are Adobe Firefly, Microsoft Designer**,** DALL-E, Midjourny, and Canva.

4.2.1 Adobe Firefly

Firefly is a logical extension of Adobe's technology developed over the last 40 years, inspired by the principle that people should be able to put their ideas into the world exactly as they envision them. The standalone web application Adobe Firefly can be accessed at firefly.adobe.com. It is safe for commercial usage and uses generative AI to greatly improve creative workflows while providing new methods to create, communicate, and brainstorm. Adobe offers the Firefly web app in addition to the larger Firefly family of creative generative AI models, as well as Firefly-powered features in Adobe Stock and its flagship applications (Get Started, 2024).

Adobe made some significant announcements regarding Adobe Firefly during Adobe MAX in October 2023. Among these is the switch to Adobe Firefly picture 2, a new version of the picture generator. The upgraded algorithm recognizes more landmarks and cultural symbols and comprehends text prompts better. It also has new Share from Firefly and Save to Library functions, as well as Prompt Guidance, which teaches users to rewrite or expand prompts (Coleman, 2023).

Jonathan Wernersson and Rickard Persson, from Jönköping University in Sweden, mentioned in their thesis titled Exploring the Potential Impact of AI on the Role of Graphic Content Creators: Benefits, Challenges, and Collaborative Opportunities (2023) that: "Firefly is a bot developed by Adobe that allows users to send prompts in the chat, which the bot then uses to generate a graphic. While similar to DALL-E 2 in some ways, Adobe Firefly provides unique 10 features and capabilities that are relevant for the research on the use of AI in the graphic design industry. Those features include typography treatments and different design modes such as "art", "photography" or "graphic". *Fig.2* represents a design made by Adobe Firefly software generated using the prompt: Pastel painting, parrot, among colorful tropical flowers, pastel painting technique, pastel texture, light blending, intricate details, soft, vibrant, layered, beautiful.

Fig. 2: An image made by Adobe Firefly software

4.2.2 Microsoft Designer

Microsoft Designer is an artificial intelligence-powered graphic design and photo editing application (Microsoft Designer, n.d). Microsoft Designer's Image Creator is powered by Open AI's most advanced image-generation model, DALL-E 3. It has the same quality findings as DALL-E, but it is free. Another significant advantage of this AI generator is that you can use it from the same location as Microsoft's AI chatbot, Copilot (previously Bing Chat). This implies that, in addition to visiting Image Creator's standalone website, you may instruct it to make photos for you in Copilot. To render a picture, simply ask Copilot to draw you the image you want (Ortiz, 2024). Microsoft Designer is now accessible as a standalone app, but it may also be accessed using Copilot in applications such as Word and

PowerPoint. With the Copilot Pro subscription, the users may generate pictures and designs in Word and PowerPoint (Mishra, 2024). *Fig.3* represents a design made by Microsoft Designer using the description: a man from Oman walking in the mart, wearing an Omani traditional outfit, holding his goat.

Fig. 3: An image made by Microsoft Designer software

4.2.3 DALL-E

DALL-E 3 is the upgraded version of DALL-E 2. What makes it one of the best free AI image generation tools is its increased power. It is designed to turn the user's words into detailed, highly realistic images (Saad, 2024).DALL-E, a 12-billion parameter version of GPT–3, was trained to produce images from text descriptions. It was discovered that it is capable of a wide range of tasks, such as anthropomorphizing animals and objects, putting disparate ideas together in logical ways, and altering already-existing images (Hanna,2023).

4.2.4 Midjourny

Midjourney is a software system that generates digital images depending on text parameters. It was built by a team led by David Holz using the high-level programming language Python (Jaruga-Rozdolska,2022). Midjourney is an AI art service that generates artwork using input prompts from a Discord channel [...] The reason Midjourney's approach can draw so precisely is that it understands the rules of painting and art and has learned from millions of art boards that have been entered into its program (Hanna,2023). Midjourney has surely gained popularity among many online groups who use the Discord platform to connect with it. It has earned a reputation as the industry leader in image generation, due to significant improvements in its capabilities in a relatively short period (Loewy, 2024).

4.2.5 Canva

One of the most widely used and easily accessible graphic design tools available is Canva, which makes it simple to generate a wide range of materials, including presentations, social media posts, posters, and brochures. It features an easy-to-use drag-and-drop interface, a big library of free photos, and pre-made designs (Garrett, 2024). Canva had more than 60 million active users worldwide by September 2021, suggesting its popularity among non-professional designers and small enterprises. Canva has found success in recent years, despite the challenges that other tools and organizations have faced (Loewy, 2024).

4.3 Future impacts of artificial intelligence technologies in graphic design

Generative AI tools are shaping the future of graphics design. Generative AI refers to AI technologies that generate visual or written content based on text input (Fatima, 2023). There is a lot of interest in how AI will affect the graphic design sector. Increased productivity is one of the main benefits since it is thought to expedite the entire design process. However, there's a chance that AI will affect design quality overall by possibly reducing human creativity and leading to more homogenized and predictable design. While AI may not be as creative as a human designer, it can be extremely adept at carrying out a human designer's vision. Although artificial intelligence (AI) might not be able to produce completely original designs, it can automate the design process and speed it up, giving designers more time to concentrate on creativity and innovation. Instead of trying to replace people, the emphasis should be on incorporating AI systems into the design process (Fatima, 2023).

Future AI applications are predicted to prioritize speed and optimization. The software's features, including data analysis and design ideas, enable designers to generate designs faster and with fewer resources. When a designer develops a design draft, artificial intelligence analyzes the data to generate the most effective design to give to the designer. It helps designers make better decisions by providing alternate designs. Rearranging data allows designers to complete projects (ÇEKEN & AKGÖZ, 2024).

V. METHODOLOGY

This study has used a quantitative research method. The quantitative approach is represented in an online questionnaire survey. Quantitative research is an effective way to acquire trustworthy and accurate quantitative data. Data is collected, evaluated, and presented numerically, therefore the results are extremely dependable and objective (Fleetwood, 2024).

4.1 Data Collection Instrument

To collect data a Survey was prepared and distributed. The questionnaire was prepared based on the literature review of the study. It was designed using Google Forms and it consists of three types of questions and they are Likert scale, checkboxes, and short answer text questions. The questionnaire was divided into Three sections which are as follows in order : (1) Demographic Details, (2) The extent to which graphic designers prefer to use AI software to accomplish design enterprises, (3)The most used AI programs by graphic designers. This survey was conducted in Arabic and English languages based on the target group so that Arabic is their mother tongue. Moreover, it facilitates answering questions for the target group and also reaches a larger number of respondents.

4.1.1 Instrument Test

A pilot test was conducted for this study as a scale to test the plan and method of research, in addition to testing the feasibility, reliability, and validity of the study design. To complete this task, a lecturer at the College of Education at Sultan Qaboos University, who specializes in graphic design tested the questionnaire. Moreover, research assistants from the Omani Studies Center were utilized to test the questionnaire and verify its validity and suitability for research purposes.

4.2 Population and Sample

This study targeted graphic designers in the age groups over 18 years both males and females. Specifically, it targeted graphic designers from Oman who work regularly in the government and private sectors, in addition to freelance designers. The number of respondents to the questionnaire reached 30 from various educational levels, starting from high school to doctorate.

4.3 Data Analysis

The data collected was examined through descriptive statistics. Descriptive analysis transforms raw data into information that is simple to analyze and understand. It rearranges, orders, and manipulates data to produce short summaries that demonstrate what occurred during a study. Descriptive statistics summarize and explain the characteristics of a given data set by providing brief descriptions of the sample and data measurements (Hayes,2024).

VI. FINDINGS AND DISCUSSIONS

Most of the respondents to the questionnaire were female, with the percentage reaching 80%, which is more than three-quarters of the respondents. Moreover, the majority of the respondents belong to the age group (25-34) constituted 60%, while the rest of them lie in the (34-44) age bracket. This shows the modernity of the field of artificial intelligence and the category that keeps up with it. Most of the respondents have bachelor's degrees, representing 83.3%. While respondents who have a diploma or its equivalent come in second place with a percentage of 10%. Respondents holding Masters and Doctorate degrees have obtained the same percentage, which is 3.3%. Regarding the employment status of respondents, the percentage of full-time employed as graphic designers was exactly 66.7%. While the percentage of respondents who were freelance designers was approximately 33%.

5.1 To what extent do graphic designers rely on AI software to accomplish their design enterprises?

Fig 4 represents bar charts of the results which show the extent to which graphic designers prefer to use AI software to accomplish design enterprises. About 50 % of the participants indicated that they sometimes use artificial intelligence software (AI) in different tasks. The statement is general, and it does not specify whether the AI programs belong to the field of graphic design. This is to estimate how familiar graphic designers in Oman are with AI software in general, regardless of their field. According to a Pew Research Center study of 11,004 people in the United States conducted December 12-18, 2022, 27% say they use AI at least many times a day, while another 28% say they connect with it once a day or several times each week. According to this self-reported statistic, 44% believe they do not regularly use AI software (Kennedy et al., 2023). The same percentage which is 50% was obtained from the second statement which is: "I use artificial intelligence software when I design" and most respondents chose the option: "sometimes". Around 36% of the designers sometimes prefer using software with artificial intelligence tools instead of software which is without artificial intelligence tools while only 20% always prefer that, and they constitute the second-highest percentage. About 40% of the respondents sometimes entirely rely on artificial intelligence to complete a task in graphic design, while only one respondent answered with "always" approximately 3%. However, the low percentage is considered a positive indicator that the designer does not rely entirely on artificial intelligence programs, and thus he gives space to show some of his creativity in the work. Nicole Steinberg (2023) a former graphic designer mentioned in her article titled "Why designers should embrace AI (and not fear it)" that:

People have a unique ability to think creatively and generate novel ideas. While AI may develop designs using specified parameters, it cannot compete with human designers' creativity and originality.

The results show that About 30% of the respondents are always up-to-date on the latest updates in artificial intelligence programs and tools in graphic design, while only one of the respondents indicated that he is not at all aware of the latest developments in artificial intelligence. Keeping up with the latest developments allows people to stay ahead of the curve and spot possibilities for innovation.

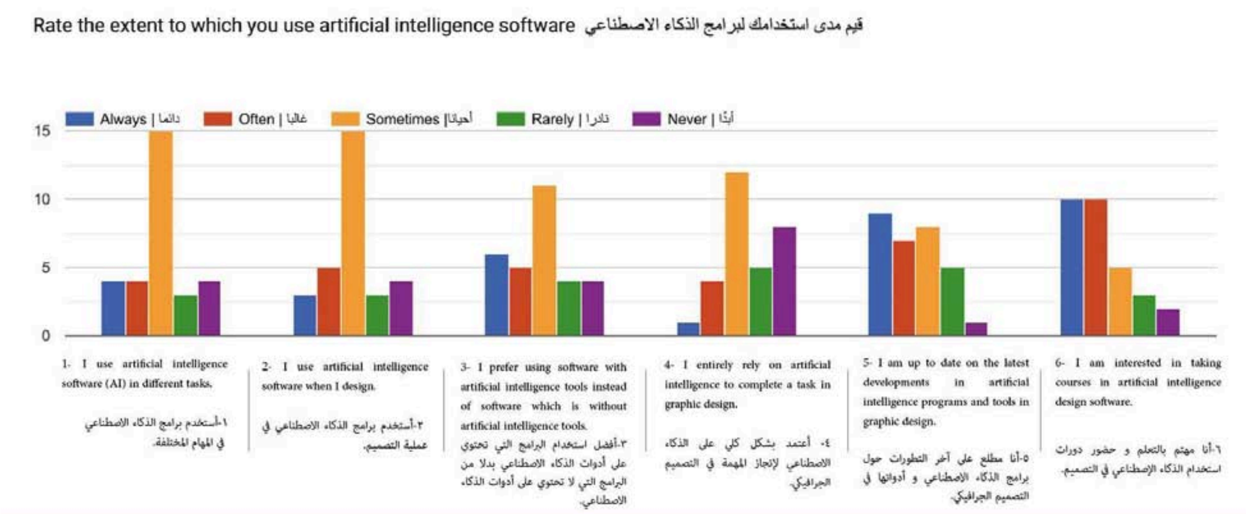


Fig. 4: Bar charts of the results which show the extent to which graphic designers prefer to use AI software

As for the statement "I am interested in taking courses in artificial intelligence design software", about 33% of the respondents responded with "always" and those who responded with "often" got a similar ratio which is 33%. These high percentages demonstrate the extent to which designers are susceptible to learning and acquiring knowledge in the field of artificial intelligence.

5.2 What are the most used AI software by graphic designers in Oman?

This section consists of five questions, three of which allow the respondent to answer more than one choice per question. The last two questions require the respondent to give short answers.

(A) The software in the questions were chosen due to their global popularity among users in general and designers in particular. The programs mentioned in the questions as options to choose from are: Adobe Firefly, Microsoft Designer, DALL-E, Midjourney, and Canva. There are also two more options, which are: All of the above and None of the above. Table 1 represents the results of three (choose multiple) questions with the results.

Table 1: The extent of the respondent's knowledge and use of specific AI programs

Questions	Number of the Responses out of 30						
	Adobe Firefly	Microdoft Designer	DALL-E	Midjourney	Canva	All of the above	None of the above
1- Which of the following AI programs have you already heard of?	15	7	2	7	20	5	2
2- Which of the following AI software have you already tried ?	13	4	3	9	15	-	4
3- Which of the following AI programs do you usually use?	8	3	1	4	11	-	9

- From *table 1*, it is clear that Canva software has the highest reputation among graphic designers in Oman compared to other software, while Adobe Firefly software ranks second.
- The results indicate that designers have tried using the following AI programs more than others, and they are, in order from the most to the least, as follows: Canva, Adobe Firefly, Midjoyrney, Microsoft Designer, and DALL-While there are about 4 designers who have not tried any of the mentioned software.
- According to the results, most of the designers usually and regularly use Canva software in their design works, while a large number of them do not rely on this specific software to do their daily work in design.

The results indicate that there is still some reservation from some designers about using programs that contain artificial intelligence tools for several reasons, the most important of which is that there are more professional software that preceded these software that give real results without relying on machines to accomplish tasks. The other reason is the lack of quality of some artificial intelligence programs in giving results, so they do not meet the expectations of the designer. In addition, most programs still do not recognize some cultures in generating images.

(B) The software mentioned in *table 2* are based on the respondent's opinion and no specific options have been set in the questionnaire. The respondent was asked to write a short answer about his favorite AI software and then the reason why he preferred this specific software.

Table. 2: Favorite AI software according to survey respondents.

Suggested software	Number of Respondents	Reasons for choosing this particular software

Adobe Firefly	2	- *easy to use* - *easy to understand*
Microsoft designer	2	- easy to use - meets design needs
Midjourny	4	- accurate in extracting data - high quality and gives amazing results. - It gives more expectations than I expected in terms of creativity and inspiration. - It facilitates the process of extracting the idea, saves time, and is sometimes effective. - Helps generate ideas.
Canva	8	- Easy to use - provides different templates (variety in templates in different fields)
Photoshop	3	- Get used to using this software before the advent of AI tools
Illustrator	1	- Get used to using this software before the advent of AI tools
Vidnoz AI	1	- It has most of the design tools I need.
Leonardo	1	- Helps generate ideas.
Nothing	7	- I don't use AI software , I don't have enough knowledge

The results in *table 2* indicate that the largest number of respondents prefer to use Canva software due to its ease of use and also because of the availability of various design templates. Canva's ease of use leads to high user satisfaction, making it a valuable tool for proper design. Canva promotes user creativity, which is an essential component of graphic design (Bimantoro & Fitriarti, 2024).

By the way, the designers who do not prefer to use AI programs are seven out of thirty designers with different educational levels between diploma, bachelor's, and doctorate. Only two of them are freelance designers while the rest work full-time. It is clear here that most of the designers who work full-time are bachelor's degree holders and this category is usually specialized in the field and has an academic certificate in design. This means that they have studied and practiced using professional design programs before adding AI tools. Therefore, it is likely that they will not accept new, less professional programs that may not give them the expected results. While AI design tools are capable of comprehending and reproducing the more conventional rules, their comprehension of context and subtleties is still lacking (Vettorino, 2023). Applications like Canva are essential in helping people without specialized training to create graphics. However, a grasp and implementation of fundamental design principles are still required to produce a powerful and distinctive visual identity. This demonstrates that even with Canva's value as a tool, graphic designers must still adhere to fundamental design principles to produce truly original work with a distinctive visual style. AI designers could not have the same originality and intuition as human designers, even though they can follow rules exactly every time. They are unable to produce original and inventive designs.

VII. CONCLUSION

In summary, Keeping up with modernity and accepting change requires a period that is not short. This study has proven that several designers are still not confident in using artificial intelligence programs in their work. The reasons may be due to a lack of awareness of using these tools in a correct way that increases their efficiency instead of considering them as tools that may replace their jobs. Artificial intelligence technology is constantly evolving, leading to the creation of new technologies and algorithms that facilitate computer-aided design. These advancements not only increase the efficiency and quality of the design process but also indicate the direction in which artificial intelligence technology is headed. While some tasks in graphic design may be automated by AI, it is unlikely to ever fully replace human creativity and knowledge. While AI can improve productivity and optimize processes, human interaction is still essential for creativity, problem-solving, and quality assurance. As we move forward, the harmonic cooperation of human intellect with artificial intelligence will continue to push the boundaries of design innovation. Future research could look into the long-term effects and future evolution of user practices as AI tools become more integrated into design workflows.

VIII. RECOMMENDATIONS

Based on foregoing findings and conclusions, the following recommendations were advanced :

1. Including the subject of artificial intelligence in design within the academic study of graphic design in universities. This is due to its great importance in keeping pace with the achievements of institutions working in this modern field of knowledge and also opening up to what has been achieved.
2. Providing finance support for full-time designers by their institutions to engage in AI-related training courses in graphic design.
3. Providing a database for independent designers in Oman and facilitating their joining training courses in the field of artificial intelligence in design. In addition, guiding them to the necessity of using artificial intelligence to keep pace with modernity.

REFERENCES

1. Ali Elfa, M. A., & Dawood, M. E. (2023). Using artificial intelligence for enhancing human creativity. Journal of Art, Design, and Music, 2(2). https://doi.org/10.55554/2785-9649.1017
2. Bimantoro, W., & Fitriarti, E. A. (2024). "CANVA", is a graphic design application that does not have a visual design identity in fundamental concepts of design. At-Tadbir: Jurnal Ilmiah Manajemen, 123–133. https://doi.org/10.31602/piuk.v0i0.15601
3. Çeken, B., & Akgöz, B. (2024). The impact of artificial intelligence on design: The example of dall-e. Sanat ve Tasarım Dergisi, 14(1), 374–397. https://doi.org/10.20488/sanattasarim.1506116
4. Coleman, K. (2023, October 14). Adobe Firefly: Everything you need to know about the AI image generator. Creative Bloq. https://www.creativebloq.com/features/everything-you-need-to-know-about-adobe-firefly.
5. Deng, J., & Chen, X. (2021). Research on Artificial Intelligence Interaction in Computer-aided Arts and Crafts. Mobile Information Systems, 2021, 1–14. https://doi.org/10.1155/2021/5519257
6. Fleetwood, D. (2024, April 16). Quantitative research: What it is, practices & methods. QuestionPro.https://www.questionpro.com/blog/quantitative-research/#What_are_the_Advantages_of_Quantitative_Research

7. Garrett, K. (2024, July 11). 12 AI tools for graphic design to test out in 2024. Filestage. https://filestage.io/blog/ai-tools-for-graphic-design/
8. Get started with Adobe Firefly. (2024, March 26). Adobe Help Center. https://helpx.adobe.com/firefly/using/firefly-overview.html
9. Hanna, D. (2023). The use of Artificial Intelligence Art Generator "Midjourney" in artistic and advertising creativity. Journal of Design Sciences and Applied Arts, 4(2), 42–58. https://doi.org/10.21608/jdsaa.2023.169144.1231
10. Hayes, A. (2024, June 27). Descriptive statistics: Definition, Overview, types, and examples. Investopedia. https://www.investopedia.com/terms/d/descriptive_statistics.asp
11. How experienced designers of enterprise applications engage AI as a design material. CHI Conference on Human Factors in Computing Systems, 1–13. https://doi.org/10.1145/3491102.3517491
12. How to analyze data in 7 simple steps: A definitive guide. (2024, August 16). Indeed. https://www.indeed.com/career-advice/career-development/analyzing-data
13. Huo, H., & Wang, F. (2022). A study of artificial intelligence-based poster layout design in Visual Communication. Scientific Programming, 2022, 1–9. https://doi.org/10.1155/2022/1191073
14. Jaruga-Rozdolska, A. (2022). Artificial Intelligence as part of future practices in the architect's work: Midjourney Generative Tool as part of a process of creating an architectural form. Architectus, 3(71), 95–104. https://doi.org/10.37190/arc220310
15. Janjetović, L., Velić, T., & Popa, M. (2023). XII INTERNATIONAL CONFERENCE ON SOCIAL AND TECHNOLOGICAL DEVELOPMENT. In XII INTERNATIONAL CONFERENCE ON SOCIAL AND TECHNOLOGICAL DEVELOPMENT PROCEEDINGS (pp. 371–376). Trebinje; University PIM.
16. Kennedy, B., Tyson, A., & Saks, E. (2023, February 15). Public awareness of artificial intelligence in everyday activities. Pew Research Center. https://www.pewresearch.org/science/2023/02/15/public-awareness-of-artificial-intelligence-in-everyday-activities/
17. Loewy, S. (2024, January 12). Artificial Intelligence Design Tool Statistics & Trends in 2023. Marq. https://www.marq.com/blog/artificial-intelligence-design-tool-statistics-trends-in-2023
18. Meron, Y. (2022). Graphic design and artificial intelligence: Interdisciplinary challenges for designers in the search for research collaboration, in Lockton, D., Lenzi, S., Hekkert, P., Oak, A., Sádaba, J., Lloyd, P. (eds.), DRS2022: Bilbao, 25 June - 3 July, Bilbao, Spain. https://doi.org/10.21606/drs.2022.157
19. Microsoft designer. (n.d.). Microsoft 365. https://www.microsoft.com/en-in/microsoft-365/microsoft-designer
20. Mishra, P. (2024, July 18). Microsoft's AI-powered designer app is now available on IOS and Android. Business Standard. https://www.business-standard.com/technology/tech-news/microsoft-s-ai-powered-designer-app-is-now-available-on-ios-and-android-124071800521_1.html
21. Munzier, D. Z., Akbar, M., Sonni, A. F., & Irwanto, I. (2024). Challenges faced by graphic designers in developing brand communication within the realm of Artificial Intelligence. Advances in Social Science, Education and Humanities Research, 639–649. https://doi.org/10.2991/978-2-38476-236-1_67
22. Mustafa, B. (2023). The Impact of Artificial Intelligence on the Graphic Design Industry. *Social Science Journal*, *13*, 243–255.
23. Ortiz, S. (2024b, June 21). The best AI image generators to try right now. ZDNET. https://www.zdnet.com/article/best-ai-image-generator/
24. Regassa Hunde, B., & Debebe Woldeyohannes, A. (2022). Future prospects of computer-aided design (CAD) – A review from the perspective of Artificial Intelligence (AI), extended reality, and 3D printing. Results in Engineering, 14, 1–9. https://doi.org/10.1016/j.rineng.2022.100478

25. Rivas, D. (2024, August 27). 17 best examples of how AI is already used in our everyday life. Digital Media Ninja. https://digitalmedianinja.com/blog/17-best-examples-of-how- ai-is-already-used-in-our-everyday-life
26. Saad, M. (2024, May 17). أفضل 6 أدوات مجانية لتوليد الصور بالذكاء الاصطناعي في 2024: بقلم: محمد سعد: موقع نثري. نثري.https://acesse.dev/snT4G
27. Sindhura, S. P., & Abdul, A. (2021). Virtues and shortcomings of artificial intelligence in graphic design arena. International Journal Advanced Research Engineering and Technology (IJARET), 12(3), 825–833.
28. Steinberg, N. (2023, March 29). Why designers should embrace AI (and not fear it). Stryve Digital Marketing. https://www.stryvemarketing.com/blog/why-designers-should-embrace-ai-not-fear-it/
29. Vettorino, M. Z. (2023, November 28). Why AI won't replace web designers [+examples of why]. Hubspot. September 9, 2024. https://blog.hubspot.com/website/will-ai-replace-web-designers
30. Wernersson, J., & Persson, R. (2023). Exploring the potential impact of AI on the role of graphic content creators: Benefits, challenges, and collaborative opportunities (thesis).
31. Yildirim, N., Kass, A., Tung, T., Upton, C., Costello, D., Giusti, R., Lacin, S., Lovic, S., O'Neill, J. M., Meehan, R. O., Ó Loideáin, E., Pini, A., Corcoran, M., Hayes, J., Cahalane, D. J., Shivhare, G., Castoro, L., Caruso, G., Oh, C., ... Zimmerman, J. (2022).

This page is intentionally left blank

Achieving Path Reversal in OTN by Control Plane Programmability

Vishal Vasant Deshpande & Dr. Shripad Prabhakar Mohani

ABSTRACT

SDN paradigm has been successfully used in IP networks to bring agility and programmability. Applying SDN to OTN (Optical Transport Networks) has some challenges due to circuit switched nature and physical characteristics of OTN. If SDN and OTN can be modified to adapt to each other, there are immense benefits that can be derived from this combination. This paper presents an innovative design by modifying components in OTN and incorporating SDN concepts along-with it. This design encompasses key SDN concepts such as control plane and data-plane separation, central view of network domain for decision making and software-based programmability. This design uniquely implements the flow-reversal in Optical Ring by programmability that its control plane offers. In traditional OTN, it is not only time consuming but also a complex process to obtain this type of flow reversal as it requires hardware configurations. The design presented here manages to rapidly configure the flow using its centralized control-plane programmatically. The design was implemented and thoroughly tested using simulator tool OptiSystem. The design is generic and does not depend on any specific vendor component. The design is flexible and can co-exist in end-to-end network setup. This design aims to help fellow researchers in their research work related to SDON (Software Defined Optical Network) and can be used as-is or with suitable modifications. This paper is organized into logical sections, starting with Introduction section that gives background and problem description.

Keywords: OTN (Optical Transport Network), optical control plane, SDON (software defined optical Network), optiSystem, openflow, SDN (software defined network).

Classification: ACM Code: C.2.1

Language: English

Great Britain Journals Press

LJP Copyright ID: 975817
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science & Technology

Volume 24 | Issue 2 | Compilation 1.0

Achieving Path Reversal in OTN by Control Plane Programmability

Vishal Vasant Deshpande[α] & Dr. Shripad Prabhakar Mohani[σ]

ABSTRACT

SDN paradigm has been successfully used in IP networks to bring agility and programmability. Applying SDN to OTN (Optical Transport Networks) has some challenges due to circuit switched nature and physical characteristics of OTN. If SDN and OTN can be modified to adapt to each other, there are immense benefits that can be derived from this combination. This paper presents an innovative design by modifying components in OTN and incorporating SDN concepts along-with it. This design encompasses key SDN concepts such as control plane and data-plane separation, central view of network domain for decision making and software-based programmability. This design uniquely implements the flow-reversal in Optical Ring by programmability that its control plane offers. In traditional OTN, it is not only time consuming but also a complex process to obtain this type of flow reversal as it requires hardware configurations. The design presented here manages to rapidly configure the flow using its centralized control-plane programmatically. The design was implemented and thoroughly tested using simulator tool OptiSystem. The design is generic and does not depend on any specific vendor component. The design is flexible and can co-exist in end-to-end network setup. This design aims to help fellow researchers in their research work related to SDON (Software Defined Optical Network) and can be used as-is or with suitable modifications. This paper is organized into logical sections, starting with Introduction section that gives background and problem description. The next section is Literature Review which covers review of relevant studies and gaps. Then Methodology section briefly covers research methodology used, how results are collected and stored etc. The Design section depicts the solution by explaining high-level design approach, reasons behind design choices and the detailed design. The Result section denotes the observed outcome and confirms that it matches with expected results. Finally, the conclusion section summarizes key points, findings, contribution and future scope.

Keywords: OTN (Optical Transport Network), optical control plane, SDON (software defined optical Network), optiSystem, openflow, SDN (software defined network).

Author α: Dept. of E&TC, *Govt. College of Engg, Jalgaon* Jalgaon, Maharashtra.
σ: Head Of Dept. E&TC, *Govt. College of Engg, Jalgaon* Jalgaon, Maharashtra.

I. INTRODUCTION

Global mobile data traffic forecast by ITU, indicates data traffic to grow at an annual growth rate of around 55% in 2020-2030. Bandwidth demand is rapidly increasing by up-to 4x per year (predominantly between data centres). As per Cisco statistics [9], the fastest growing component of data-centre traffic is Global cloud traffic. It is seen growing by a huge 40% annual growth rate. Today's cloud, mobility and content services offered by service providers have dynamic needs. In order to meet those needs, Telecom operators require more dynamic and programmable network infrastructure with high capacity. It is essential to support rapid service deployment and provide real-time responsiveness to capacity changes.

Optical Transport Networks (OTN) promises the high capacity and reliable bandwidth fulfilment. But they are not flexible to dynamic needs as they are static and mostly manually provisioned. Optical networks are operated with wavelengths fixed in place and not designed to dynamically change. So, they are not capable of adapting to rapid changes to deliver the flexible services like cloud. Modern Network traffic consists of short bursty traffic, as well as very high bandwidth, high-duration data flows that continue for minutes. Common examples of such high bandwidth persistent traffic are VM migrations, data migrations, or Data warehousing function (MapReduce). [10]. On the other hand, OTN is traditionally a circuit switched network and the Path is set at initial design time. Changes to path requires long time and network can't adapt to dynamically to varying traffic conditions mentioned above.

Software Defined Networks (SDN) paradigm promises flexibility and programmability in network operations. SDN was originally conceived for packet-based IP networks. Hence it is difficult to apply SDN to OTN as is. SDN achieves the programmability by decoupling the data plane and the control plane. In most of the OTN these planes are currently vertically integrated and inseparably hosted along-with data-plane. Even if some designs attempt separately hosted optical control plane, it only allows network management software interactions [16]. It does not expose its services to operator directly i.e. no API or interfaces exposed for issuing commands to control plane. Hence optical control planes cannot be programmed dynamically in present OTN.

This paper presents a design that attempts separation of optical control plane from data plane in OTN. The paper also explains the programmability of optical control plane through software instructions or configurations. The design helps control plane to get centralized view that enables it to take decisions at network domain level (instead of at network each element level). The uniqueness of design in achieving flow control (path reversal) programmability in OTN ring.

Traditionally, the design approach for IP-plus-Optical network has been, to place all the network functions within the IP layer (routing, signalling, protection). Such design uses static optical trunks interconnecting these IP layer devices. The network controller which controls IP domains does not have capabilities to configure OTN. It treats OTN merely as a fixed pipe carrying data. This paper takes a unique approach in designing OTN Ring to perform flow-control by programming OTN. Moreover, this design ensures adherence to key concepts of SDN (such as control plane separation, centralized view of control plane, software-based programmability etc), it does not mandate OpenFlow or another southbound interface (SBI). OpenFlow is not yet a complete standard, it is still undergoing significant changes [3] to adapt to OTN. Hence it can be incorporated in this design as a future scope.

II. LITERATURE REVIEW

A detailed analysis of multiple articles was carried out in regards to specific architectures or models proposed by other researchers. Abhinava Sadasivarao and others have proposed [1] programmable architecture that tries to integrate with the deployment of SDN within the Data Centres. It abstracts a transport node into a programmable virtual switch and leverages the OpenFlow protocol for control. It can be extended to packet-switched transport architectures including MPLS. But the the time involved in setting up the path using SDN Controller showed a high latency (between 5s to 7s). Also, this architecture has implicit mode where the SDN controller has a view of only the edge nodes.

Cheng, Xu and Zang from China mobile in their paper [2] have explained the requirements for advanced network architecture of Packet Transport Network. It has good focus on fast provisioning, end-to-end multi-domain and multi-layer network view. Although, there is no consideration given to optical transport networks.

S. K. N. Rao in his white-paper [3] has provided a survey of SDN and its use cases explaining benefits and limitations. It explains that a typical SDN architecture has the network intelligence logically centralized in software-based controllers. This enables the control logic to be designed and executed on a global network view. This was a key takeaway from this whitepaper. But the architecture described in it was just high level block-diagram. Moreover, it only described few options without specific recommendations, hence it was found to be closer to study-paper.

R. Vilalta et. al. in their paper [4] proposed a network control architecture for multidomain and multivendor network which is organized in layers (Refer fig.1) The abstract network layer and the control-specific layer, results in a mesh of generic SDN controllers. This uses generalized multiprotocol label switching (GMPLS) protocols as their east/west interfaces. This architecture was tested for its performance on service provisioning latency and control plane overhead. As a limitation, the authors have clarified that application of this architecture are confined and scoped to a single or reduced number of operators with peering agreements. Verizon [5] has deployed similar hierarchical architecture for moving to 100G Packet-optical transport network.

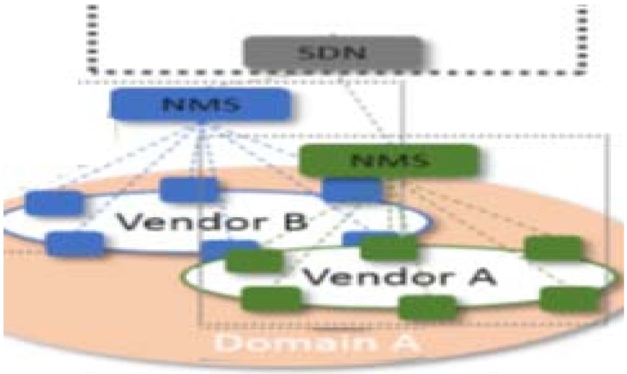


Fig. 1: Hierarchical SDN Architecture for OTN

Industry body Open Networking Forum (ONF) has also proposed [6] an architecture with a parent "super" controller. The network orchestrator will abstract the details of the optical transport layer. It also enables end-to-end provisioning of services and provides open interfaces to client SDN applications. It is possible to have multiple technology controllers per domain, with this architecture.

Optical network model based architectures claim to provide open and vendor agnostic management of optical equipment. Thomas Szyrkowiec et. al. have investigated [7] Optical Network Models and their application to SDN Management. They have surveyed and compared important optical network models. They proposed an intent interface for creating virtual topologies which is integrated in the existing model ecosystem. This is an interesting architectural approach. It makes it easy to achieve software-based control but requires the network topology virtualization. In paper [10], a similar virtual modelling using VOLTHA was described. VOLTHA is emerging as a standard for virtualization of OLT (online termination). The modelling approaches in [7] and [10] has limitations, as assumptions like every cross-connection between input and output ports is possible, does not hold for a realistic optical network model. Also on a physical network level, the analogue nature leads to network constraints, which can't be modelled accurately.

A simplified architecture is proposed and analysed by RK Jha and Burhan NML in their paper [8]. In this architecture, control of ethernet elements (routers) and optical ring is performed by a single SDN controller. Moreover, the architecture includes standardized hardware for OpenFlow switches and standard interface for OpenFlow communication. Communication between ethernet switches and optical network devices is setup using OEO converter. Communication between SDN controller and optical network devices is setup using OpenFlow agents. Hence this architecture will need enhancements before applying to all-optical network. It can also be noted that the optical hardware (ROADM), based on banyan architecture switch, combines the data plane along with switching control inside it.

III. METHODOLOGY

The experimental design methodology was used to implement the design, run tests, collect the results and demonstrate outcomes. The type of experimental design chosen was Absolute Experiment. The input was given (in the form of WDM lambdas) and output was observed by changing the control bits (or configurations). The output was recorded in the form of signal graphs and compared with expected result. Controlled variables of this experiment included input signal wavelength, OADM's drop channel setting, selection of Optical cross-connect ports. Factors not controlled include signal characteristics (like attenuation, dispersion, noise) which were not influencing factors.

This experiment was carried out in a high-end Simulator called OptiSystem. (Design section 4.2 gives more information on choice of this simulator tool.) The results were captured in the form of signal graphs shown on optical spectrum analyser (OSA). Multiple runs (iterations) were executed that gave consistent output that matches with expected output. To decide number of test-runs by using same input and to decide the variation of inputs for additional test-runs, techniques similar to sampling techniques were used. E.g. sample size of 5 test-runs with same input is considered. non-probability sampling is used for choice of input lambdas as these are in the least attenuation region (1550nm) wavelength. Judgement sampling is used for deciding add and drop channels for alternate test runs. (Note: Even though the term 'sample' is used here, it actually refers to 'test-run' or simulator iteration).

All the observations collected from these test-runs forms the empirical data collection by experiment. Simple comparison method was used to analyze the test run results. First result was compared with the theoretically expected result of the design. Following results were compared with the first result to check if any deviation. The data storage was done inside the simulator tool using the file system.

IV. DESIGN

Before proceeding to detailed design description, it is imperative to a indicate the high-level design aspects like topology and components. The type of OTN that is considered here is long-haul transport deployment e.g. between major cities. This is a very common deployment of OTN and generalizes other types (e.g. DCN = Data Centre network) also well. Bidirectional Ring topology was considered, since most of the OTN deployments follow Ring topology [18].

The key components in any ring are OADM (Optical Add-Drop multiplexer) and OXC (Optical Cross connect or Optical Switch). Some vertically integrated devices contain OADM, OXC, transponders etc in the same physical chassis. They perform forwarding (data-plane) and switching (control-plane) activities together inseparably. Hence as a design choice, such integrated devices were avoided. When the same signal was to be replicated on other fiber (or other direction), simple fork was utilized. Its possible to replace it with switch if any setup so requires. Since the experimental setup and outcomes are not influenced by physical characteristics like noise, BER, dispersion, there is no specific assumption on type of fiber or repeaters or amplifiers etc. The Ring input is in the form of WDM signal with four channels viz 193.1 THz, 193.2 THz, 193.3 THz and 193.4 THz. The drop and add channels of each OADMs are as shown in Table.1:

Table 1: OADM channel configuration

OADM	Drop Channel	Add Channel
A	193.1 THz	193.7 THz
B	193.2 THz	193.8 THz
C	193.3 THz	193.9 THz
D	193.4 THz	194.0 THz

4.1 Detailed Design

The key design considerations include:

- Data-plane and control-plane separation: Data plane focuses only on data-forwarding.
- Controller having central view instead of one node (and neighbouring nodes): This helps in deciding network domain level functions like path traversed in the ring.
- Controller issuing software instructions (configurations or scripted commands) to bring programmability, allowing change of flow dynamically.

To accomplish these key considerations, following layers are envisaged. (Refer Fig.2) The bottom layer is Data-plane, formed by four OADMs. On top of that resides a Control plane. It consists of OXC (optical switches) which are software-controlled by the controller. Controller has a full network domain view. It issues software commands to switches in control plane and accordingly the switches control the path flown through the data-plane.

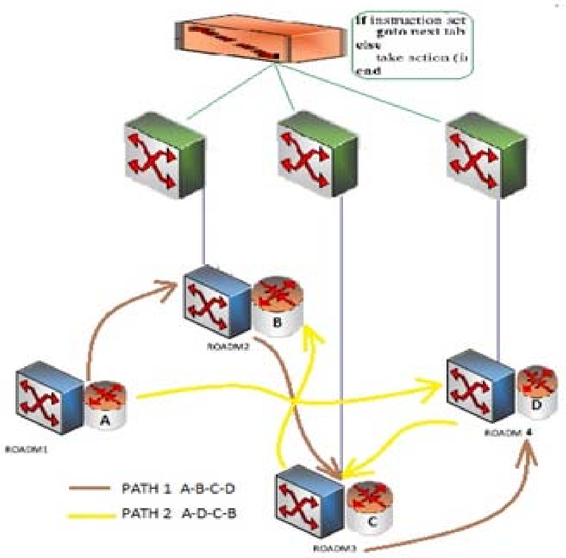


Fig. 2: Design showing control and data planes

The three OXCs are co-located i.e. placed nearby the OADMs B, C and D. But they are not hosted on or tightly couples with OADMs, like in traditional OTN. Based on instruction from controller, the OXC selects different ports and the signal coming from OADM is put onto one of the ports selected. One port allows the flow in one direction (A-B-C-D) while other port enables a reverse flow direction (A-D-C-B).

This allows to change the flow direction dynamically based on controllers input to switches in control-plane. As can be seen, OADM has no participation in flow-control. It just processes (drops a channel and adds new channel) irrespective of flow direction. This design can be used by researchers as a starting point in their SDON (Software Defined Optical Network) research, or for training others about SDN, or for integration testing in larger network setup. This design is implemented and tested using the simulator environment and output graphs are collected.

4.2 Simulator Selection

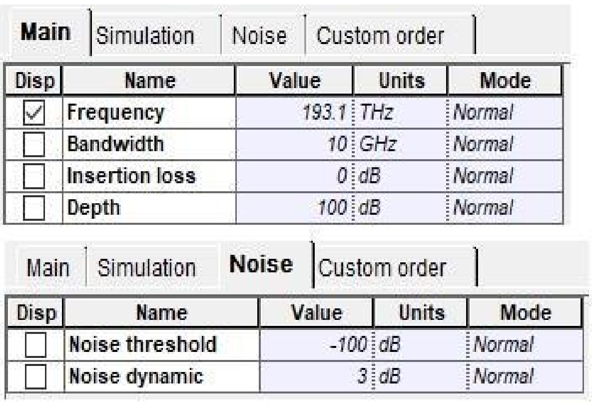
Simulation are technology tools and they help with the unreal and real-life entities to be modelled on to the computer and run under certain predefined conditions [11][12] Network simulators are widely used by the research community to evaluate new theories and hypotheses. [13] Selection of simulator can impact the outcomes, hence sufficient analysis is required before selection, especially in optical network research. [14] The author has analysed many simulators and has come up with detailed approach to select simulator for research in OTN. Using a simulator selection tool that the author himself created for ranking the simulators, OptiSystem simulator was ranked highest for implementing this design, testing it and collecting the results/measurements of test-runs.

OptiSystem is specialized simulator for optical communication systems and includes all layers of Optical Transport Network [15]. It has comprehensive library of components. It offers high model accuracy, wide range of modulation formats, powerful simulation environment and hence good fit for research purpose. It has a truly hierarchical definition of components and systems including Optical sub-system. As per author's self-experience, this tool is quite easy to install and run. The GUI is intuitive and user-friendly and provides ability to quickly design optical Networks.

4.3 Environment Setup

Latest OptiSystem (v21.0) is used for implementing this design and setting up the test-run.

- The global parameters related to optical characteristics (such as attenuation, insertion loss, power levels) were maintained as default. (as shown in Fig. 3)
- The global parameters related to simulator model are evaluated. Some minor modifications done to bit rate, padding by no of leading/lagging zeros, central frequency (as shown in Fig. 4)
- Execution related parameters are changed as: set iterations to 2 within every run. Accordingly set the signal buffer value same as iterations (=2). Set location to store the results and filename to end with the test-run number, timestamp.



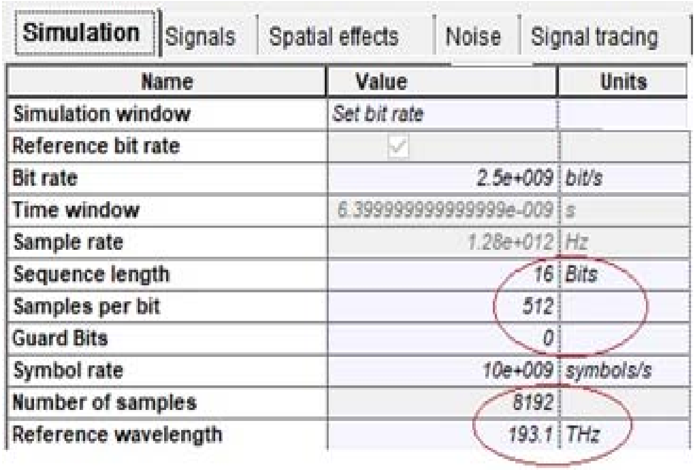
Main | Simulation | Noise | Custom order

Disp	Name	Value	Units	Mode
☑	Frequency	193.1	THz	Normal
☐	Bandwidth	10	GHz	Normal
☐	Insertion loss	0	dB	Normal
☐	Depth	100	dB	Normal

Main | Simulation | Noise | Custom order

Disp	Name	Value	Units	Mode
☐	Noise threshold	-100	dB	Normal
☐	Noise dynamic	3	dB	Normal

Fig. 3: Global parameters: Optical characteristics



Name	Value	Units
Simulation window	Set bit rate	
Reference bit rate	✓	
Bit rate	2.5e+009	bit/s
Time window	6.399999999999999e-009	s
Sample rate	1.28e+012	Hz
Sequence length	16	Bits
Samples per bit	512	
Guard Bits	0	
Symbol rate	10e+009	symbols/s
Number of samples	8192	
Reference wavelength	193.1	THz

Fig. 4: Global parameters: Simulation Model characteristics

V. RESULTS

There are two iterations carried out to test run the two flows: First flow is OADM A-to-B-to-C-to-D and second flow is reverse, OADM A-to-D-to-C-to-B. Both flows worked as designed and successfully showed the designed output.

Results are collected as the graphs from Optical Spectrum Analyzer (OSA). Depending upon the test, the OSA is placed at various points to traverse the flow of signal step-by-step and check the correctness of channels for that step. The resulting graphs at OADM A, B, C and D are shown below for both the test runs. These results are stored in flat file. The file names end with test-run, iteration number, timestamp information. OADM A is a common Ring ingress point for both the flows. In both the test-runs, same input signal (WDM) is given to Ring at OADM-A as shown in fig. 5.

Flow 1 Output: Output taken at OADM-A, B, C and D shows each OADM drops a channel and adds a channel on its input signal. So, at output of D (i.e. Ring output), as expected the original four channels are dropped and new four channels are added, as seen in fig 6 (OADM A & B) and Fig 7 (C & D).

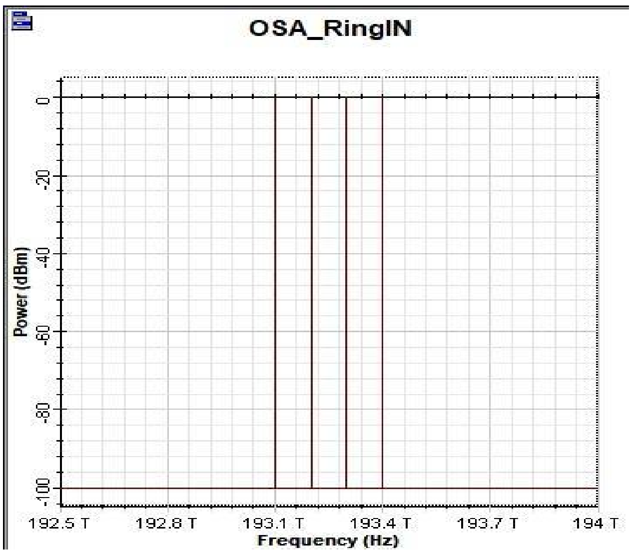


Fig. 5: Ring Input (WDM) at OADM-A

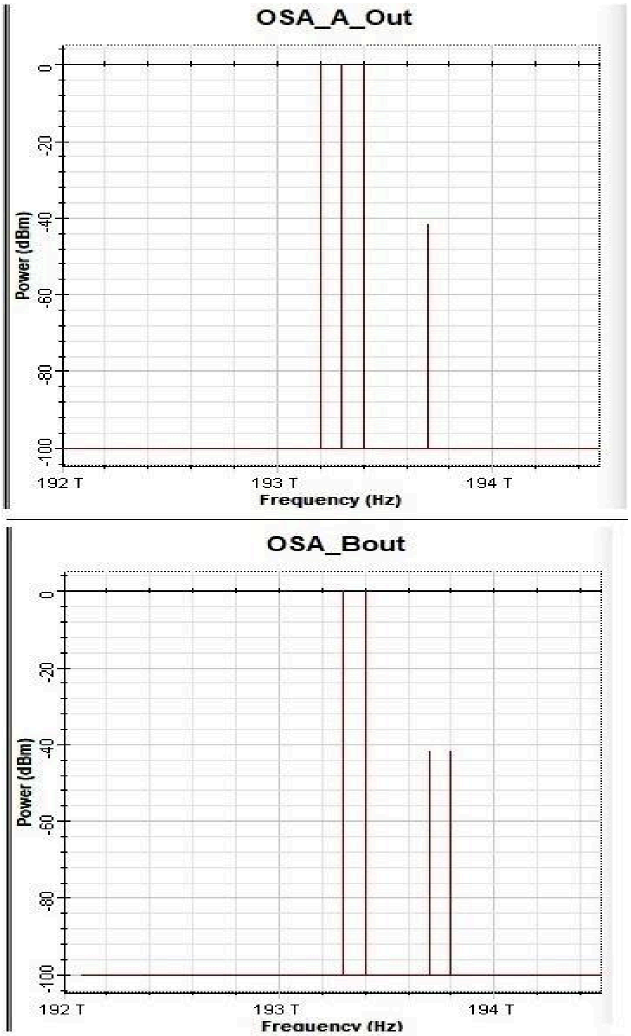


Fig. 6: Output at OADM-A & B for flow 1

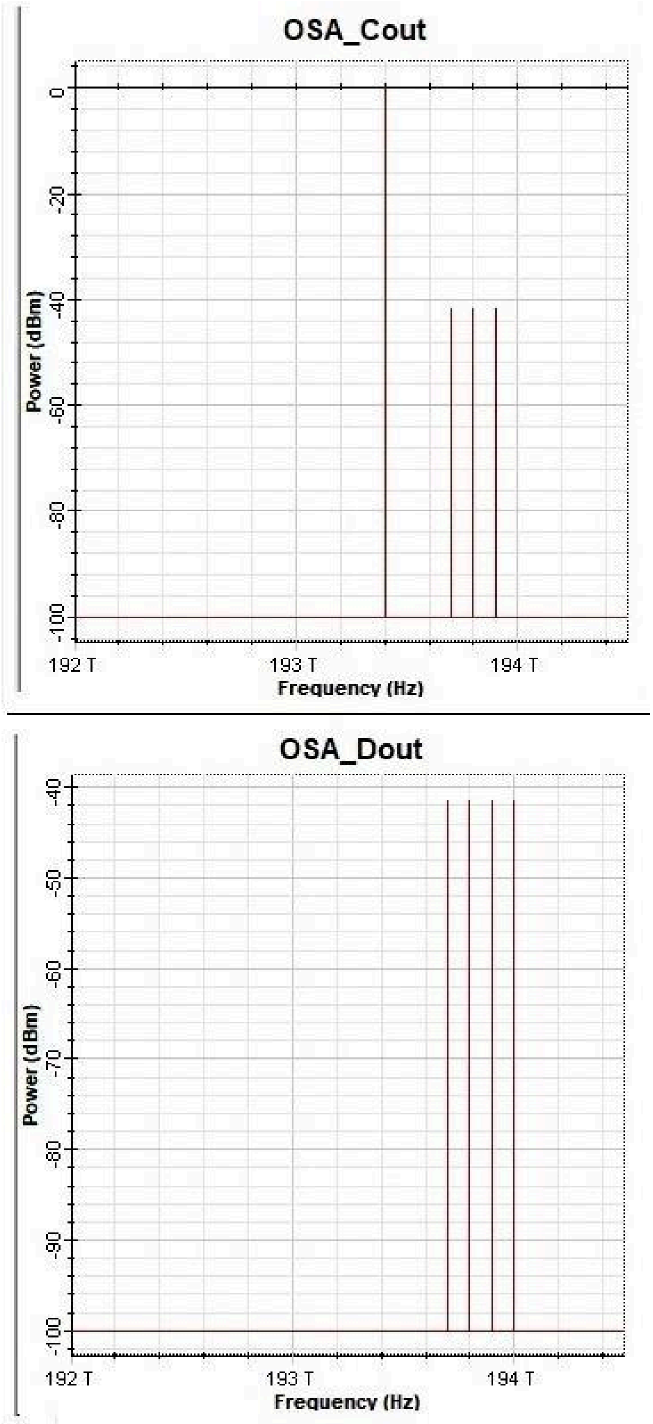


Fig. 7: Output at OADM-C & D for flow 1

In the reverse flow test-run, output of OADM-A is same (as same input is provided). Thereon, there is a difference in flow. Output at B is now Ring output due to flow reversal and B shows only new four channels added (as expected for reverse flow). Output of OADM A, D show in Fig 8 and output of OADM C & B is shown in Fig 9.

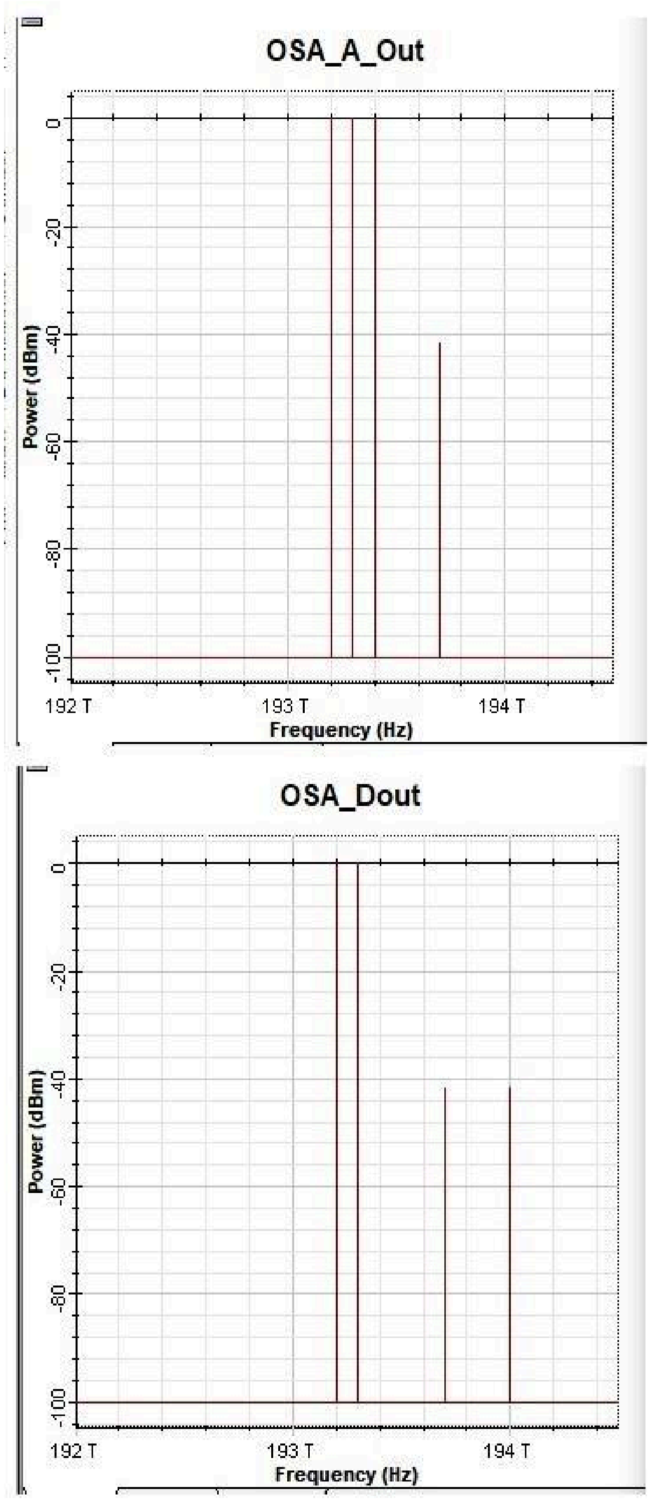


Fig. 8: Output at OADM-A & D for flow 2

This confirms that flow-reversal design has worked as desired. The controller changes the configurations and sends instruction (bit-sequence) to switches to reverse the flow. This shows that the programmability of flow in the OTN can be achieved by control-plane.

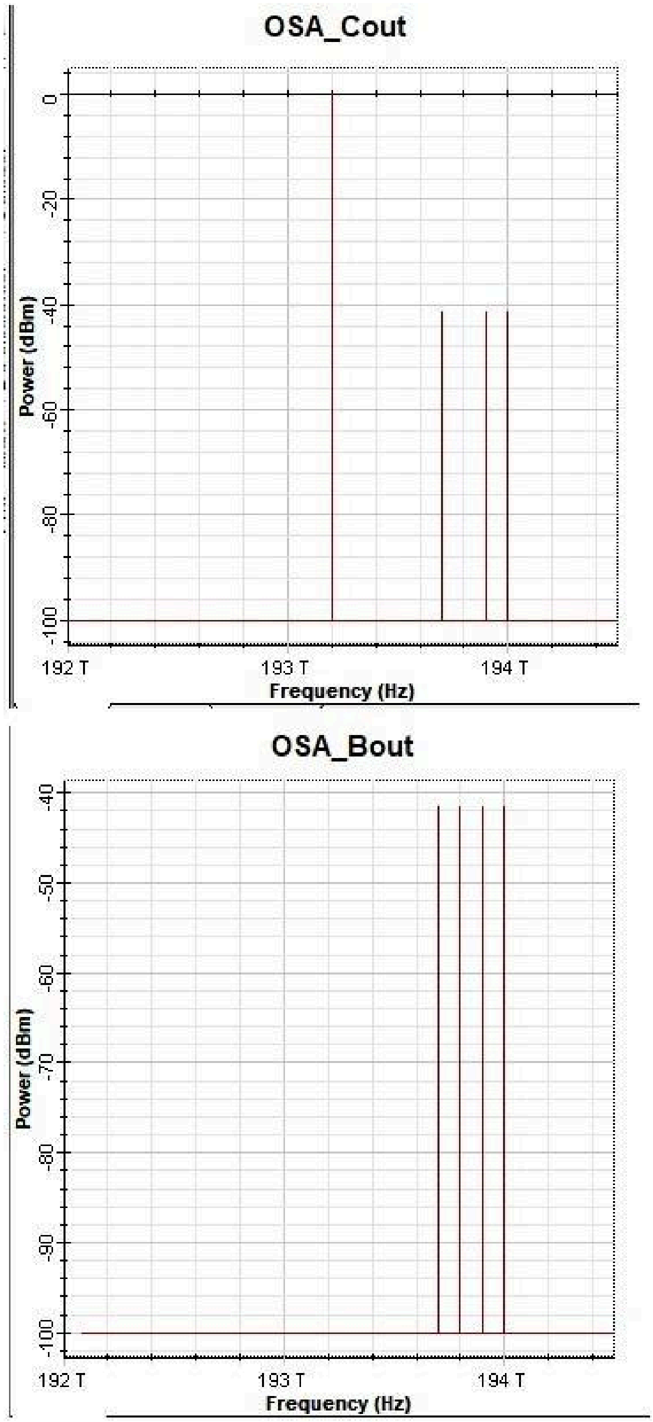


Fig.8: Output at OADM-C & B for flow 2

VI. CONCLUSION

It was experimentally shown that the proposed design is capable of achieving flow-control in OTN with the help of programmability offered by its central control plane. The design when implemented works as desired and flow reversal in Ring topology was observed. This unique design follows key concepts in SDN (like control-data plane separation, software-based programmability, central view of network domain etc). Hence it can be inferred that it is possible to apply SDN concepts to OTN, although SDN can't be applied to OTN in the same way it is applied to IP networks. SDN expects optical networks to be modified to adapt to SDN e.g. by clear separation of data-plane and control-plane. This paper presents a design with optical control-plane and data-plane logically separated. Flow control is dynamically achieved by programmability offered by central control-plane that has full view of optical ring network. Thus, the paper has illustrated SDN-like programmability in OTN using flow-control scenario. Researchers can reuse this design in their SDON research work as-is or by modifying as per their needs, the design is flexible. Academician can use it for training SDON concepts or for integration testing with another network setup. This tested, working and flexible to modify network design acts as a reliable starting point. With it, researchers can make alterations with confidence and refer back to this reliable stratum. It can also be concluded that OptiSystem simulator was suitable choice for implementing and testing this OTN design. As a future work, OpenFlow can be added to this design for SBI (southbound interface)

VII. ACKNOWLEDGMENT

The author wishes to thank OptiWave OptiSystem [15] for allowing temporary access to licensed copy of the latest version (v21.0) of OptiSystem Simulator.

REFERENCES

1. Abhinava Sadasivarao et. al, Bursting Data Between Data Centers: Case for Transport SDN, IEEE 21st Annual Symposium on High-Performance Interconnects, pp87-90
2. Weiqiang Cheng, ACTN Use-cases for Packet Transport Network, China Mobile, IETF, Jul 2014
3. S. K. N. Rao, SDN and its Use-Cases - NV and NFV: A State-of-the-Art survey, presented at NTIL 2014 event
4. R. Vilalta et. al, Control and Orchestration of Multidomain Optical Networks With GMPLS as Inter-SDN Controller Communication, Journal of Optical Communications and Networking Vol. 7, Issue 11, pp. B46-B54
5. Verizon readies its metro for next-generation P-OTS, Gazettabyte, June 2014
6. ONF-TR 538 : Use Cases for Carrier Grade SDN
7. Thomas Szyrkowiec, Achim Autenrieth and Wolfgang Kellerer,Optical Network Models and Their Application to Software-Defined Network Management, International Journal of Optics, Volume 2017
8. Rakesh Kumar Jha and Burhan Num Mina Llah, SDON: proposed architecture and comparative analysis, Journal of the European Optical Society-Rapid Publications (2019)
9. Cisco Global Cloud Index: Forecast and Methodology, 2015–2020, White Paper
10. Calient White Paper, “The Software Defined Hybrid Packet Optical Datacenter Network”
11. J. Heidemann, K. Mills, S. Kumar, Expanding Confidence in Network Simulation, IEEE Net-work Magazine, Vol. 15, No. 5, Apr. 2000, pp. 58- 63.
12. Siraj, Saba, A. Gupta, and Rinku Badgujar. Network simulation tools survey, International Journal of Advanced Research in Computer and Communication Engineering Vol. 1, Issue 4, pp. 199-206, 2012.
13. Atta ur Rehman Khana, Sardar M. Bilalb, Mazliza Othmana, A Performance Comparison of Network Simulators for Wireless Networks, IEEE International Conference on Control System, Computing and Engineering, Malaysia, 2012, pp. 34-38
14. Vedran Miletić, Branko Mikac, Matija Dzanko, Modelling optical network components: A network simulator-based approach, ResearchGate Conference Paper, October 2012
15. Optisystem product website https://optiwave.com/products/optisystem
16. Yao Li and D. Kilper, Optical Physical Layer SDN, Journal of Optical Communication Networks, Vol-10, pp110-121

17. Shu Namiki et. al, Dynamic Optical Path Network, The 12th International Conference on Optical Internet Proceedings
18. Raimena Viesllari et. al., Scalability Analysis of SDN-controlled Optical Ring MAN with Hybrid Traffic, IEEE ICC 2014 - Optical Networks and Systems, pp3283-3288.

This page is intentionally left blank

Security Through Elliptic Curves for Wireless Network in Mobile Devices

Dr. E. Kesavulu Reddy

Sri Venkateswara University

ABSTRACT

The basic principle is "A function is easy to evaluate but its invert is infeasible unless a secret key is known".

It is mathematically proved that security of cryptographic algorithms does not imply its implementation of security of the system against Side-channel Attacks. The security of the system lies in the difficulty of extracting k from P and Q. It is essential to secure the implementation of cryptosystems in embedded devices against side-channel attacks. These attacks monitor the power consumption or the Electromagnetic emanations of a device Ex smart cards or mobile devices. The attacker's goal is to retrieve partial or full information about a long-term key that is employed in several ECSM executions.

We are implementing a secret key to avoid retrieving the valuable information by the attacker through Simple Power Analysis Attacks.

Keywords: elliptic curve cryptography, simple power analysis, differential power analysis.

Classification: ACM Code: C.2.0

Language: English

Great Britain Journals Press

LJP Copyright ID: 975825
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science & Technology

Volume 24 | Issue 2 | Compilation 1.0

Security Through Elliptic Curves for Wireless Network in Mobile Devices

Dr. E. Kesavulu Reddy

Professor_Department of Computer Science Sri Venkateswara University Tirupati, AndhraPradesh India.

ABSTRACT

The basic principle is "A function is easy to evaluate but its invert is infeasible unless a secret key is known".

It is mathematically proved that security of cryptographic algorithms does not imply its implementation of security of the system against Side-channel Attacks. The security of the system lies in the difficulty of extracting k from P and Q. It is essential to secure the implementation of cryptosystems in embedded devices against side-channel attacks. These attacks monitor the power consumption or the Electromagnetic emanations of a device Ex smart cards or mobile devices. The attacker's goal is to retrieve partial or full information about a long-term key that is employed in several ECSM executions.

We are implementing a secret key to avoid retrieving the valuable information by the attacker through Simple Power Analysis Attacks.

Keywords: elliptic curve cryptography, simple power analysis, differential power analysis.

I. INTRODUCTION

Elliptic curve cryptosystems (ECCs) are suitable for implementation on devices with limited memory and computational capability such as smart cards and also with limited power such as wireless handheld devices. This is due to the fact that elliptic curves over large finite fields provide the same security level as other cryptosystems such as RSA for much smaller key sizes.

Considering power analysis attacks, there are two main types that were presented by Kocher et al. These are simple and differential power analysis attacks (referred to as SPA and DPA respectively). Both of them are based on monitoring the power consumption of a cryptographic token while executing an algorithm that manipulates the secret key.

The traces of the measured power are then analyzed to Hence, DPA attacks are, in general, more powerful than the SPA attack. Randomization of the data processed at some instant is essential in resisting this type of attack.Electromagnetic emanations present another powerful side channel since the information is leaked from the device via more than one channel and is a function of space as well as of time. In [2], Agrawal et al. presented both simple and differential electromagnetic analysis attacks on smart cards and on a Palm pilot . In these attacks they conclude that software countermeasures rely on signal information reduction, which is achieved by "randomization and/or frequent key refreshing within the computation", which agrees with the concept of resisting DPA attacks.

The point addition operation consists of finite field operations carried in the underlying field K. We denote the field inversion by I, the multiplication by M, the squaring by S. The point addition is denoted by A. When the two operands of the addition are the same point, the operation is referred to as point doubling and is denoted by D.

II. WINDOW METHODS

This method is sometimes referred to as the m–ary method. There are different versions of window methods . What is common among them is that, if the window width is w, some multiples

London Journal of Research in Computer Science & Technology

of the point P up to $(2^w - 1)$P are precomputed and stored and k is processed w bits at a time. k is recoded to the radix 2^w. k can be recoded in a way so that the average density of the nonzero digits in the recoding is $1/(w + \xi)$, where $0 \le \xi \le 2$ depends on the algorithm. Let the number of precomputed points be t, in the Precomputation stage, each point requires either a doubling or an addition to be computed also depending on the algorithm.

This ECSM method is suitable for unknown or fixed-point P. The cost is Storage: t points, where $2^{w-2} \le t \le 2^{w-1}$ depending on the algorithm.
Precomputation: t point operations (A or D).
Expected running time:

$$(n-1)\,D + n\frac{n}{w+\xi}\,A,$$

where $0 \le \xi \le 2$ depending on the algorithm. Note that the number of doublings is between n – w and n – 1.

2.1. Simultaneous multiple point multiplication

This method is used to compute kP +lS where P may be a known point. This algorithm was referred to as Shamir's trick in [10]. If k and l are n-bit integers, then their binary representations are written in a 2 × n matrix called the exponent array. Given width w, the values iP +jS are calculated for $0 \le i, j < 2^w$. Now the algorithm performs d = $\lceil n/w \rceil$ iterations. In every iteration, the accumulator point is doubled w times and the current 2 × w window over the exponent array determines the precomputed point that is to be added to the accumulator.

Algorithm 2.1. Simultaneous multiple point multiplication *(Shamir-Strauss method)*

Input: Window width w, d = $\lceil n/w \rceil$,

k = $(K^{d-1}, \ldots, K^1, K^0)_{2^w}$ l= $(L^{d-1}, \ldots, L^1, L^0)_{2^w}$, and P, S ∈ $E(F^q)$. 3 Also according to [Ber01], it is originally due to Straus [Str64].

Output: kP + lS.

1. Precomputation. Compute iP + jS for all i, j ∈ $[0, 2^w - 1]$.
2. Q ← $K_{d-1}P + L_{d-1}S$.
3. for i from d – 2 down to 0 do
3.1 Q ← 2^wQ.
3.2 Q ← Q + $(K_iP + L_iS)$.
4. Return(Q).

Storage: $2^{2w} - 2$ points. For w = 1, 3 points.
For w = 2, 15 points.
Precomputation:

$(2^{2(w-1)} - 2^{w-1})\,D + (3 \cdot 2^{2(w-1)} - 2^{w-1} - 1)\,A$.
For w = 1, 1 A.
For w = 2, 1 D + 11 A.

Expected running time: $(d-1)w\,D + \left(\frac{(2^{2w}-1}{2^{2w}}\,d - 1\right)A$.

For w = 1, $(n-1)\,D + (\frac{3}{4}n - 1)\,A$.

For w = 2, $(n-1)\,D + (\frac{15}{32}n - 1)\,A$.

Using sliding windows can save about 1 4 of the precomputed points and decrease the number of additions to $\frac{n}{w+(1/3)}$, which is about 9% saving for w ∈ {2, 3}.

2.2. Interleaving method

This method is also a multiple point multiplication method, that is we want to compute $\sum k^j P_j$ for points P^j and integers k^j. In the comb and simultaneous multiplication methods, each of the precomputed values is a sum of the multiples of the input points. In the interleaving method, each precomputed value is simply a multiple of one of the input points. Hence, the required storage and the number of point additions at the precomputation phase is decreased at the expense of the number of point additions in the main loop. This method is flexible in that each k^j can have a different representation, e.g., different window size, as if a separate execution of a window method is performed for each $k^j P^j$ with the doubling step performed jointly on a common accumulator, as shown in [12]. As an illustration, we provide the following algorithm that computes kP + lS where both k and l are represented to the same base 2^w.

Algorithm 2.2. Interleaving method

Input: width w, d = $\lceil n/w \rceil$,

k = $(K_{d-1}, \ldots, K_1, K_0)_{2^w}$ l= $(L_{d-1}, \ldots, L_1, L_0)_{2^w}$, and P, S $\in$ $E(F_q)$.

Output: kP + lS.

1. Precomputation. Compute iP and iS for all i $\in$ $[0, 2^w - 1]$.
2. Q $\leftarrow$ K_{d-1}P.
3. Q $\leftarrow$ Q L_{d-1}S.
4. for i from d – 2 down to 0 do

4.1 Q $\leftarrow$ 2^wQ.

4.2 Q $\leftarrow$ Q + K_i P.

4.3 Q $\leftarrow$ Q + L_i S.

5. Return(Q).

Storage: $2^{w+1} - 2$ points.

Precomputation: 2(w-1) D + 2(2^w – w – 1) A.
Expected running time: w(d – 1) D + (2d – 1).

$$\frac{(2^{2w}-1)}{2^{2w}} A$$

In general, if different basis and/or representations are used for k and l, we have

Storage: 2t points, where $2^{w-2} \le t \le 2^{w-1}$ depending on the particular window algorithm used as discussed in Section 2.1.3.

Precomputation: 2t point operations (A or D).

Expected running time: (n – 1) D + $2\frac{n}{w+i}$ A, where $1 \le i \le 2$ depending on the algorithm

2.3. SPA Attack on ECCs and its Countermeasures

Coron [3] has transferred the power analysis attacks to ECCs and has shown that an unaware implementation of EC operations can easily be exploited to mount an SPA attack. Moreover, it may also enable it to recognize the exact instruction that has been executed. For example, if the difference in power consumption between point doubling (D) and point addition (A) is obvious in their respective power traces, then, by investigating one power trace of a complete execution of a double-and-add algorithm, the bits of the scalar k are revealed. That is, whenever a D is followed by A, the corresponding bit is k_i = 1, otherwise if D is followed by another D, then k_i = 0. This sequence of point operations is referred to as the DA sequence.

Window methods process the key on a digit (window) level. The basic version of this method, that is where _ = 0 in Section 2.1.3, is inherently uniform since in most iterations, w D operations are followed by 1 A, except for possibly when the digit is 0. Therefore, fixed-sequence window methods were proposed in order to recode the digits of the key such that the digit set does not include 0.

2.4. DPA Attack on ECCs and its Countermeasures

As for the SPA attack, Kocher et al. were the first to introduce the DPA attack on a smart card implementation of DES. Techniques to strengthen the attack and a theoretical basis for it were presented by Messerges et al. in [3]. Coron applied the DPA attack to ECCs [3].

In order to resist DPA attacks, it is important to randomize the value of the long-term key involved in the ECSM across the different executions. Some of the countermeasures that were based on randomizing the key representation were proven to be inadequate since the intermediate point computed in the accumulator Q at certain iteration remained one of two possible values . The constancy of the value of this intermediate point is an integral part in the success of first-order DPA attacks.

A potential DPA countermeasure is known as key splitting . It is based on randomly splitting the key into two parts such that each part is different in every ECSM execution. An additive splitting using subtraction is attributed to Clavier and Joye . It is based on computing

$$kP = (k - r)P + rP, \qquad (I)$$

The authors mention that the idea of splitting the data was abstracted in [5]. where r is a n-bit random integer, that is, of the same bit length as k. Alternatively, Ciet and Joye [8] suggest the following additive splitting using division, that is, k is written as

$$k = \lfloor k/r \rfloor + (k \bmod r). \qquad (1)$$

Hence, if we let $k_1 = (k \bmod r)$, k_2 2 = $\lfloor k/r \rfloor$ and $S = rP$, we can compute

$$KP = k_1 p + k_2 P \qquad (II)$$

where the bit length of r is n/2. They also suggest that (II) should be evaluated with the Shamir-Strauss method as in Algorithm 2.3. However, they did not mention whether the same algorithm should be used to evaluate (I). The following multiplicative splitting was proposed by Trichina and Bellezza [0] where r is a random integer invertible modulo u, the order of P. The scalar multiplication kP is then evaluated as

$$kP = [kr^{-1} \pmod u)] (rp) \qquad (III)$$

To evaluate (III), two scalar multiplications are needed; first R = rP is computed, then kr^{-1} R is computed.

III. KEY SPLITTING METHODS

3.1. Introduction

We discuss different forms of key splitting along with their strengths and weaknesses. We also discuss the candidate SPA-resistant algorithms and compare the resulting performance when combined with each form of key splitting. At the end of the chapter, we present countermeasures to DPA attacks on the ECDSA and the ECMQV algorithms.

This approach was suggested by Clavier and Joye in [CJ01] and revisited by Ciet [Cie03] as follows. In order to compute the point kP, the n-bit key k is written as

$$k = k_1 + k_2 ,$$

such that $k_1 = k - r$ and $k_2 = r$, where r is a random integer of length n bits. Then kP is computed as

$$kP = k_1 P + k_2 P. \qquad (4.1)$$

It is important to note that each of the terms should be evaluated separately and their results combined at the end using point addition. That is the multiple-point multiplication methods that use a common *accumulator to save doubling operations*. Whether at the bit level (w ≡ 1) or window level (w > 1)-should not be used, even when a countermeasure against SPA is employed. This observation is based on the following lemma.

Let $_{b \to a}$ denote $\lfloor k(\bmod 2^{b+1}) 2^a \rfloor$ or, simply, the bits of k from bit position b down to bit position a, with b ≥ a.

Lemma 3.2 Let splitting scheme I at the end of some iteration j, 0 < j ≤ n − 1, there are only two possible values for Q, those are $[k_{n-I-j}]$ P or $[k_{n-I-j} - 1]$ P.

Proof. Algorithm 2.3—and similarly Algorithm 2.4—computes the required point by scanning

$$k_1 = (k_{1\,n-1}, \ldots, k_{1\,0})2 \text{ and}$$

$k_2 = (k_{2\,n-1}, \ldots, k_{2\,0})_2$ from the most significant end down to the least significant end. Hence, at the end of iteration j, the accumulator Q contains the value

$$Q = k_{1\,n-I-j} P + k_{2\,n-I-j} P \qquad (4.2)$$

$$= [k_{1\,n-I-j} + k_{2\,n-I-j}] P.$$

We can write k, k_1 and k_2 as

$$k = k^{n-I-j^j}\, 2^j + k^{j-1-0}$$

$$k^i = k^{i\ n-I-j^j}\, 2^j + k^{i\ j-1-0} \qquad (4.3)$$

Since $k = k^1 + k^2$ we have

$$k^{1\ j-I-0} + k^{2\ j-I-0}\, j = k^{j-I-0} + b\, 2^j \text{ where } b \in \{0, 1\} \qquad (4.5)$$

and

$$k^{1\ n-I\to j} + k^{2\ n-I\to j} = k^{n-I\to j} - b$$

The DPA attack would proceed in the same way, whether the algorithm processes a single bit or a digit per iteration, though it would be more involved in the latter case depending on the digit size. The attacker can double the number of traces gathered and compute the necessary intermediate points as if there was no countermeasure in place.

3.2 Modular Division:

In the following algorithm, a and b are integers internally represented each by an array of w-bit digits. The length of each array is $d = \lceil n/w \rceil$ digits. Note that for the modular inversion, as mentioned by Savas and Ko¸c [25], b needs not be less than the modulus u, but be in $[1, 2^m - 1]$, where m = dw. Also note that the values R^2 mod u, where $R = 2^m$, and u' are computed once per modulus, i.e., per curve.

Algorithm 3.4. Modular division

Input: u: a n-bit prime, $d = \lceil n/w \rceil$, m = dw, R^2 (mod u) = $(2m)^2$ (mod u), u' = u^{-1} mod 2^w, a $\in$ [1, p – 1] and b $\in$ $[1, 2^m - 1]$.

Output: ab^{-1} (mod u).

1. Compute b^{-1} R (mod u) using Algorithm 6.6.

2. Compute x = $a(b^{-1} R)R^{-1}$ (mod u) using Algorithm 6.5.

3. Return(x).

The following algorithm is Algorithm 14.36 in [21]. We include it here for the sake of completeness.

Algorithm 3.5Montgomery multiplication

Input: u: a n-bit prime, $d = \lceil n/w \rceil$, m = dw, u' = u^{-1} mod 2^w, x = $(x_{d-1} \ldots, x0)2^w$ and $y = (y_{d-1} \ldots y0)2^w$.

Output: $xy2^{-m}$ (mod u).

1. A ← 0. // A = (ad, a_{d-1}, . . . , a0)2w

2. for i from 0 to d – 1 do

2.1 ui ← $(a_0 + x_i y_0) \bmod 2^w$

2.2 A ← $(A + x_i y_0 + u_i m)/ 2^w$

3. if (A > u) then

A A ← u.

4. Return(A).

The following algorithm was presented by Savas and Ko¸c as the modified Kaliski-Montgomery Inverse.

Algorithm 3.6. Montgomery inversion

Input: u: a n-bit prime, $d = \lceil n/w \rceil$, m = dw, R^2 (mod u) = $(2^m)^2$ (mod u), $u' = u^{-1} \bmod 2^w$ and b $\in$ $[1, 2^m - 1]$.

Output: b^{-1} R (mod u).

1. Compute f and x = b^{-1} 2f (mod u) using Algorithm 6.7, Where n ≤ f ≤ m + n.

2. if (f ≤ m) then

2.1 x ← $xR^2 R^{-1}$(mod u) using Algorithm 6.5. // x = $b^{-1} 2^{m+f}$ (mod u)

2.2 f f + m. // f > m, x = b–12f (mod u)

3. x ← $x2^{2m-f} R^{-1}$ (mod u) using Algorithm 6.5.

// $x = b^{-1} 2^f\, 2^{2m-f}\, 2^{-m} = b^{-1} 2^m$ (mod u)

4. Return(x).

IV. EXISTING SYSTEM

Nevine Maurice Ebied modified the almost Montgomery inverse algorithm of [ScKK00] to be resistant to SPA attacks as in the following algorithm. Swap Address(c, d) denotes interchanging the memory addresses of the integer's c and d. This is an inexpensive operation, hence its usage as a dummy operation to balance the branches of the main loop. We implemented the "if" statement in steps 3.4 and 3.5 such that the number of conditions checked per loop iteration is always three. In assembly language, this can be easily ensured. Written in Java, step 3.4 is implemented as

If ((xLSb = = 0) && (xLSb = = 0) && (xLSb = = 0)).

If the condition is false, due to short-circuit evaluation, the flow control will move to the following "if" after the first check, otherwise, it will perform the check three times. The following "if"—step 3.5—is similar but with the condition checked only two times if((yLSb = = 0) && (yLSb = = 0)).

Algorithm 3.7. Almost Montgomery inverse :Input: u: a n-bit prime,

d = $\lceil n/w \rceil$, m = dw and b ∈ [1, 2^m – 1].

Output: f and $b^{-1} 2^f$ (mod u), where n ≤ f ≤ m + n.

1. x ← u; y ← b; r ← 0; s ← 1.

2. f ← 0.

3. while (v > 0) do

3.1 U ← x – y; V ← –U.

3.2 T ← r + s.

3.3 f ← f + 1.

3.4 if (((lsb (x)=0))) then
// This "if" is special SwapAddress(x, U);
SwapAddress (x, U) // dummy
SHR(x); SHL(s).
3.5 else if ((lsb(y) = 0)) then
// This "if" is special SwapAddress(y, V);
SwapAddress
(y, V) // dummy
SHR(y); SHL(r).
3.6 else if (V >= 0) then
SwapAddress(y, V);
SwapAddress(s, T)
SHR(y); SHL(r).
3.7 else
SwapAddress(x, U);
SwapAddress(r, T)
SHR(x); SHL(s).

4. T ← u – r; V ← u + T .

5. if (T > 0) then
Return(f, T)
else
Return(f, V).

The drawback of this algorithm is that an SPA of the number of iterations of the main loop directly leaks the value of f. If f is uniformly distributed, the search space of b is reduced from 2^w to $2^{m - \log 2^m}$, which is not a significant reduction. It is interesting to study how f is actually distributed.

4.1. PROPOSED SYSTEM

We modified the Nevine Maurice Ebied's Almost Montgomery inverse and A SECRET KEY of [ScKK00] to be resistant to SPA attacks as in the following algorithm.

Algorithm:4.2. E.KESAVULU REDDY (EKR) Modified Montgomery Inversion

Input: u: a n-bit prime, d = $\lceil n/w \rceil$,

m = dw, R^2 (mod u) = $(2^m)^2$ (mod u), u = u^{-1} mod 2^w and b ∈ [1, 2^m – 1], t is Secret key.

t: No of precomputed points 1≤ t ≤n

W: Window width least significant of bit

$$2^{w-z} \quad t \leq 2^w - 1$$

Output: b^{-1} R (mod u).

Select a number b such that $(b, 2^m) = 1$

Compute b^{-1} such that $bb^{-1} \equiv 1 (\mathrm{mod}\ 2^m)$

3. If f > m then x = $b^{-1} 2^f$ (mod u)
⌧ x = $b^{-1} 2^f$ (mod u)
4. If f ≤ m then
5. x ← $R^2 R^{-1} (\mathrm{mod}\ u)$ ⌧ R = 2^m
6. x = $b^{-1} 2^{m+f}$ (mod u)
7. f ← m+f
8. Return(x)

In the EKR modified Montgomery Inverse Algorithm of Savas and Koc , we select f such that gcd $\left(b,2^{m}\right)=1$, m ≤ f ≤ m +n . So b is not reduced from 2^{m} to $2^{m-\log 2^{m}}$. Therefore, this is significant reduction and hence f is not uniformly distributed and it can't leak the value package javaapplication1;

```
import java.util.ArrayList;
import java.util.Scanner;
public class Main {
 public static void main(String[] args)
{
Scanner s=new Scanner(System.in);
System.out.println("Enter a Prime Integer(U):");
int u=s.nextInt();
String binU=Integer.toBinaryString(u);
int n=binU.length();
System.out.println("Enter a secrete Key(t):");
int t=s.nextInt();
double w=0;
for(int wi=1;wi<=t;wi++)
{
 int i1=(int)Math.pow(2,wi-2);
 int i2=(int)Math.pow(2,wi)-1;
 if(i1<=t && t<=i2)
 {
  w=wi;
  break;
 }
}
System.out.println("Window width (W): "+w);
int d=(int)Math.ceil(n/w);
System.out.println("Length of Each Array (d) :"+d);
int m=(int)(d*w);
System.out.println("....(m) :"+m);
int on1=(int)Math.floor(Math.pow(2,m-1));
int n2=(int)Math.pow(2,m);
for(int i=m;i<on1;i++)
{int n1=i;
while(n1!=n2)
{
 if(n1>n2)
  n1=n1-n2;
  else
  n2=n2-n1;
  }
//System.out.println("GCD of two number is "+n1+"and i is "+i);
if(n1==1){on1=i;break;}
}
System.out.println("B="+on1);
int b=on1;
int b_inverse=0;
for(int i=1;i<b;i++)
{
int t2m=(int)Math.pow(2,m);
int temp=t2m*i;
temp=temp+1;
b_inverse=temp/b;
if((temp%b)==0)
{
break;
}
}
System.out.println("B inverse: "+b_inverse);
ArrayList<Integer> af=new ArrayList<Integer>();
for(int i=n;i<=(m+n);i++)
{   af.add(i);
}
int x=0;
for(int f=n;f<=(n+m);f++)
{
if(f>m)
{
 x=(int)(b_inverse*Math.pow(2, f))/u;
System.out.println("f="+f+"\nx="+x);
}
else if (f<m)
{
 x=(int)(b_inverse*(Math.pow(2,m+f)))/u;
System.out.println("f="+af.get(f)+"\nx="+x);
}
}
}
}
```

INPUT AND OUTPUT

Enter a Prime Integer(U):
111
Enter a secrete Key(t):
117
Window width (W): 7.0
Length of Each Array (d) :1
....(m) :7
B=7
B inverse: 55
f=8
x=126
f=9

x=253
f=10
x=507
f=11
x=1014
f=12
x=2029
f=13
x=4059
f=14
x=8118
BUILD SUCCESSFUL (total time: 29 seconds)

V. CONCLUSION

We modified the Nevine Maurice Ebied's Almost Montgomery inverse and A New variant of [ScKK00] of Montgomery Inversion i.e is the EKR Modified Montgomery Algorithm to be resistant to SPA attacks . The EKR Modified Montgomery Inverse Algorithm eliminate the number of Iterations of the main loop directly leaks the value of f and also it is mathematically proved that f is uniformly distributed with a significant reduction

A function that is easy to evaluate but infeasible to invert unless the secret trapdoor t is known. So, the attacker cannot guess the key (t) to retrieve the valuable information in smart cards and mobile devices .

REFERENCES

1. Nevine Maurice Ebied's Key Randomization Counter Measures To Power Analysis Attacks On Elliptic Curve Cryptosystems Ph.D. thesis, University of Waterloo, Ontario, Canada, 2007
2. D. Agrawal, B. Archambeault, J. R. Rao & P. Rohatgi. The EM Side- Channel(s): Attacks and Assessment Methodologies. Internet Security Group, IBM Watson Research Center.ps. 2, 3
3. J.-S. Coron. "Resistance against differential power analysis for elliptic curve cryptosystems". In Cryptographic Hardware and Embedded Systems –CHES '99, LNCS, vol. 1717, pp. 292–302. Springer-Verlag, 1999. 2, 22, 24, 158, 170, 180, 181, 186
4. S. Chari, C. S. Jutla, J. R. Rao & P. Rohatgi. "Towards sound approaches to counteract power-analysis attacks." In Advances in Cryptology – CRYPTO '99, LNCS, vol. 1666, pp. 398–412. Springer-Verlag, 1999. 24
5. C. Clavier & M. Joye. "Universal exponentiation algorithm a first step towards provable SPA-resistance". In Cryptographic Hardware and Embedded Systems – CHES '01, LNCS, vol. 2162, pp. 300–308. Springer-Verlag, 2001. 4, 24, 120
6. M. Ciet, J.-J. Quisquater & F. Sica. "Preventing differential analysis in GLV elliptic curve scalar multiplication". In Cryptographic Hardware and Embedded Systems – CHES '02, LNCS, vol. 2523, pp. 540–550. Springer- Verlag, 2003. 4, 25, 126, 173, 179
7. T. ElGamal. "A public key cryptosystem and a signature scheme based on discrete logarithms". IEEE Transactions on Information Theory, 31(4):469– 472, 1985.
8. J. Ha & S. Moon. "Randomized signed-scalar multiplication of ECC to resist power attacks". In Cryptographic Hardware and Embedded Systems – CHES '02, LNCS, vol. 2523, pp. 551–563. Springer-Verlag, 2002. 3, 24, 27, 31, 40, 58, 70, 75, 93, 101, 173
9. T. S. Messerges, E. A. Dabbish & R. H. Sloan. "Investigations of power analysis attacks on smart cards". In USENIX Workshop on Smart- card Technology, pp. 151–161. May 1999. 2, 24, 172
10. B. M¨oller. "Securing elliptic curve point multiplication against side channel attacks". In International Security Conference – ISC '01, LNCS, vol. 2200, pp. 324–334. Springer-Verlag, 2001. Extended version.pdf. 23, 164,196,19:195–249, 2000. 4, 14, 15, 20, 37, 46, 50, 58, 70, 89, 90, 91, 92, 95, 110, 112.
11. N. Theriault. "SPA resistant left-to-right integer recodings". In Selected Areas in Cryptography – SAC '05, LNCS, vol. 3897, pp. 345–358. Springer- Verlag, 2006. 23,128, 133, 164).

Dr.E.Kesavulu Reddy working as a Professor in the department of Computer Science, S.V.University,

Tirupati. He Published research papers in various National and International Journals and presented papers in national and International Conferences. He was awarded National and International awards with different prestigious organizations in India. He was honored Fellow of Computer Science Research Council **(FCSRC)** from Open association of Research society (U.S.A), Global Journals, U.S.A on 31-01-2019.

This page is intentionally left blank

Scan to know paper details and author's profile

Artificial Intelligence and the Challenges of Copyright Protection Systems – in Search of Common Norms

Dr. Saliltorn Thongmeensuk, Pitchapon Jirawongsapan & Noppasin Camapaso

ABSTRACT

The integration of artificial intelligence (AI) into the production of creative works presents significant challenges to the existing frameworks of copyright law. With AI's increasing autonomy in creative processes, traditional notions of authorship and originality are under scrutiny, necessitating a reevaluation of copyright protections. The paper proposes an evolved copyright framework that distinguishes between AI-assisted and AI-autonomous creations, suggesting clear guidelines for AI authorship and recommending the establishment of a specific copyright category for AI-generated works. This category would adjust protection levels and terms to reflect the unique nature of these works while ensuring that human creativity remains incentivized. The proposed framework aims to recalibrate copyright law to ensure the protection of human creators' rights while acknowledging AI's innovative contributions. This focused approach underscores the need for copyright law to adapt to technological advances, ensuring that it continues to fulfill its role in promoting creativity and innovation.

Keywords: artificial intelligence (AI), AI-generated works, innovation, authorship, originality, copyright law.

Classification: ACM Code: K.4.1,

Language: English

Great Britain Journals Press

LJP Copyright ID: 975826
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science & Technology

Volume 24 | Issue 2 | Compilation 1.0

Artificial Intelligence and the Challenges of Copyright Protection Systems – in Search of Common Norms

Dr. Saliltorn Thongmeensuk[α], Pitchapon Jirawongsapan[σ] & Noppasin Camapaso[ρ]

ABSTRACT

The integration of artificial intelligence (AI) into the production of creative works presents significant challenges to the existing frameworks of copyright law. With AI's increasing autonomy in creative processes, traditional notions of authorship and originality are under scrutiny, necessitating a reevaluation of copyright protections. The paper proposes an evolved copyright framework that distinguishes between AI-assisted and AI-autonomous creations, suggesting clear guidelines for AI authorship and recommending the establishment of a specific copyright category for AI-generated works. This category would adjust protection levels and terms to reflect the unique nature of these works while ensuring that human creativity remains incentivized. The proposed framework aims to recalibrate copyright law to ensure the protection of human creators' rights while acknowledging AI's innovative contributions. This focused approach underscores the need for copyright law to adapt to technological advances, ensuring that it continues to fulfill its role in promoting creativity and innovation.

Keywords: artificial intelligence (AI), AI-generated works, innovation, authorship, originality, copyright law.

I. INTRODUCTION

Copyright law protects original works of authorship, including literary, musical, artistic, and other creative works. In recent years, Artificial intelligence (AI) has been increasingly used to generate these various forms of expression generally protected by copyrights, leading to questions about authorship. AI can generally be described as the capacity of a digital computer or a robot controlled by a computer to execute activities typically linked with intelligent entities.[1] The use of AI in creative works raises legal challenges that can only be addressed by an understanding of the interplay between copyright law and technological advancements.

The use of artificial intelligence to create works of art has significant implications for copyright law. In the past, computer-generated works were not typically questioned in terms of copyright ownership, as the program was seen as a tool that facilitated the creative process, much like paper and canvas. For a work to be eligible for copyright protection, it generally needs to be original and have a human author. Many countries, such as Spain and Germany, have laws that stipulate that only work created by humans are protected by copyright. However, with the latest forms of artificial intelligence, the computer program is no longer simply a tool, as it now makes many of the creative decisions without human intervention. This creates a new challenge for copyright law. This part analyses the concept of authorship and originality in different countries and suggests how this concept should be developed to ensure that the incentives for creation can still thrive in digital age.

One of the primary legal challenges in AI-generated art is determining the author of the work. When sophisticated AI technology is employed in creating a piece of art, it frequently becomes challenging to pinpoint specific human

[1] B.J. Copeland, 'Artificial intelligence', *Encyclopedia Britannica*<www.britannica.com/technology/artificial-intelligence> accessed 20 December 2023

London Journal of Research in Computer Science & Technology

authors who contributed to the expression or compilation of parts of the work, or to determine the exact portion they created or compiled.[2]

At the heart of this issue is the concept of originality. Copyright protection is traditionally predicated on the originality of the work, which implies a human author's creative expression. Yet, AI-generated works complicate this notion, as they are often the result of algorithms processing and recombining vast datasets of existing content. This raises the question of whether such works can be considered 'original' in the copyright sense and, if so, who or what should be credited as the author—the AI algorithm, its programmer, the user who initiated the creative process, or the entity owning the AI system?

Moreover, the legal and ethical implications of AI in creative domains extend beyond the bounds of authorship and originality. They encompass broader concerns regarding the impact of AI on the creative industries, the potential for AI to disrupt traditional copyright enforcement mechanisms, and the ethical considerations of recognizing non-human entities as authors or rights holders. As AI continues to transform the landscape of creative work, these considerations become increasingly pertinent, challenging copyright scholars, practitioners, and policymakers to forge new paths in copyright law that honor the dual imperatives of promoting innovation and protecting human creativity.

This paper seeks to navigate these complex waters by offering a comprehensive analysis of the current state of copyright law in the context of AI-generated works. It explores various jurisdictions' approaches to defining authorship and originality, examines potential legal frameworks for accommodating AI's role in creative processes, and proposes principles for a balanced copyright regime that recognizes the contributions of both human and artificial creators. Through this exploration, the paper aims to contribute to the ongoing dialogue on copyright law's future in an increasingly digitized and AI-driven world.

II. THE CHARACTERISTICS OF THE AI-GENERATED ARTS AND THE IMPACTS ON ART MARKET

To understand the impact of artificial Intelligence on the art market and copyrights, it is essential to take a glance at its characteristics, how it works, how it produces art, who is involved in its development, and what the impact would be on the artist and market. At the end of this section, this study will attempt to answer the controversial question of whether AI could produce work as an author without human intervention and whether that work could be considered art.

Artificial Intelligence has been integrated into several sectors, i.e. transportation, energy, healthcare, and so on, and its role is chiefly about improving productivity and social welfare – e.g. to promote 'efficiency' [3]. AI has been adopted into the creative industry, and several AI art generators – Midjourney, Dell-E, and so on – have been introduced to the market. AI art generators use sophisticated models like Graph Neural Network (GNN) to generate arts, which require extensive image data as a vital resource. To acquire the data, AI developers often scrape image data from Pinterest or Fine Art America, the third-party platforms where artists pose their works for PR purposes.[4] [5] However, Artificial Intelligence cannot understand the art on its own. Thus, AI developers must assign human agents to label the image based on their perceptions and

[2] Jani McCutcheon, 'The Vanishing Author in Computer Generated Works: A Critical Analysis of Australian Case Law' (2012) 36 Melbourne University Law Review 915, 918.

[3] Lodewijk Heylen, 'Art and Automation: The Role of the Artist in an Automated Future' (2019) <https://iscma.scm.cityu.edu.hk/openconf/modules/request.php?module=oc_program&action=view.php&id=69&file=1/69.pdf>

[4] Melissa Heikkiläarchive, 'This Artist Is Dominating AI-Generated Art. and He's Not Happy About It' (*MIT Technology Review*, 16 September 2022) <www.technologyreview.com/2022/09/16/1059598/this-artist-is-dominating-ai-generated-art-and-hes-not-happy-about-it/> accessed 20 August 2023.

[5] Andy Baio, 'Exploring 12 Million of the 2.3 Billion Images Used to Train Stable Diffusion's Image Generator' (*Waxy*, 30 August 2022) <https://waxy.org/2022/08/exploring-12-million-of-the-images-used-to-train-stable-diffusions-image-generator/> accessed 20 August 2023.

opinions. [6]. Thus, generative AI tools such as Dall-E 3 or Stable Diffusion often put a tag or label on the image, and most of the time, these tags are the artist's name[7]; the user will use the tag to command AI to generate an image based on the prompt, e.g. the user can command the AI to generate an image of an orange with Picasso style. The algorithm will search for the image that matches the prompt and produce work based on the data, which makes the technology lack originality and could harm the original artists by hindering their sales or damaging their reputation.[8]

One of the harms the AI might pose to the artists is the threat to their sales and profitability. For example, a Polish illustrator, Greg Rutkowski, is one of the artists affected by the AI art generator tool. He was initially optimistic about the technology since it might help him reach new consumers, but he later found several works of art with his name tag created by AI. He is concerned that if this tool becomes more widespread, he might not be able to find his original work with internet search engines, which might hinder his sales[9]. Furthermore, as AI art generators become extensively used in the market, they might create a considerable displacement effect as the technology replaces human artists [10]. This phenomenon is called 'excessive automation,' where the excessive use of technology might undermine labor productivity[11]. However, this might not be the case as the characteristics of the creative industry are unique.

The creative industry is different from other industries, primarily in value creation. The value of art is related to intangible components, which are often heterogeneous among individuals, like culture, emotion, moral knowledge, or human behavior.[12] These components will determine how individuals are willing to pay for the art piece based on their perceived value, e.g. one might favor painting over video, thus, willing to pay much more than the painting. However, this does not mean that painting is better than the video.[13] Moreover, the individually perceived value is dynamic; it changes according to how society defines 'art' The definition of art in the present revolves around Anthropocentrism. Anthropocentrism is the notion that humans are the only entity possessing intrinsic value; thus, only humans can create art.[14] Various experiments confirm the prevalence of Anthropocentrism; Fortuna et al.'s (2021) findings suggest that consumers value AI arts lower than human-created arts[15]. One possible reason was depicted in the literature where the authors experimented to observe how consumers value a work of art; they found that consumers often value painting based on the time and effort spent in the production process[16]. Moreover, a Work of Art tends to be highly rated if the creators are known [17] and there is a story behind it[18]. These are

[6] Carloalberto Treccani, 'How Machines See the World: Understanding Image Labelling' (2019) <www.academia.edu/44900825/How_machines_see_the_world_understanding_image_labelling>

[7] Laurie Clarke, 'When AI can make art – what does it mean for creativity?' *The Guardian* (12 November 2022). <www.theguardian.com/technology/2022/nov/12/when-ai-can-make-art-what-does-it-mean-for-creativity-dall-e-midjourney> accessed 10 September 2023.

[8] Harry H Jiang and others, 'AI Art and Its Impact on Artists', *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society* (ACM 2023) <https://dl.acm.org/doi/10.1145/3600211.3604681> accessed 3 January 2024.

[9] ibid.

[10]ibid 6.

[11]Daron Acemoglu and Pascual Restrepo, 'Artificial Intelligence, Automation, and Work' (2019) The Economics of Artificial Intelligence: An Agenda 197

[12] Antonio Daniele and Yi-Zhe Song, 'AI + Art = Human' (the 2019 AAAI/ACM Conference on AI, Ethics, and Society (AIES '19), Association for Computing Machinery, New York, NY, USA, January 2019)

[13] Bruno Frey, 'Art: The Economic Point of View' in Alan Peacock and Ilde Rizzo (eds), *Cultural Economics And Cultural Policies* (Springer, Dordrecht 1994)

[14] Michael Straeubig, 'Do Machines Produce Art? No. (A Systems-Theoretic Answer.)' (Art Machines: International Symposium on Computational Media Art, January 2019) < https://doi.org/10.13140/RG.2.2.32258.50885>

[15] Paweł Fortuna and Artur Modliński, 'A(I)Rtist or Counterfeiter? Artificial Intelligence as (d)Evaluating Factor on the Art Market' (2021) 51(3) The Journal of Arts Management, Law, and Society 188 <https://doi.org/10.1080/10632921.2021.1887032> accessed 10 September 2023.

[16] Justin Kruger, Derrick Wirtz, Leaf Van Boven, *et al.* 'The Effort Heuristic' (2004) 40(1) Journal of Experimental Social Psychology 91 <https://doi.org/10.1016/s0022-1031(03)00065-9> accessed 8 December 2023.

[17] ibid.

[18] Leslie Snapper, Cansu Oranç, Angelina Hawley-Dolan, *et al.* 'Your Kid Could Not Have Done That: Even Untutored

consistent with Tubadji et al.'s (2021) argument that consumers' willingness to pay for art does not only depend on objective economic value but also cultural economic value, e.g. the time the human producer spends on the product, the feeling they feel at that time and so on[19]. For these reasons, excessive automation by AI in the creative industry does not seem to be the case, as consumers still favor human work over robots. However, it could replace lower-productivity jobs. Moreover, as artificial Intelligence becomes more prominent in the creative industries, it will create new jobs, markets, and opportunities.

Artificial Intelligence is a concept that has been introduced for a while in the art market. Harold Cohen has been working with algorithmic art since 1968,[20] and several artists and programmers have used the technology to create art for some time[21]. However, these artists were scarce as the technology was in an infant state; the technology only possessed a limited degree of autonomy as it required humans to instruct them. However, AI now requires only minuscule human intervention; a few only require humans to intervene in development. Consequently, as automation sometimes can work independently, the new question arises about who would be the author of the work created by AI. Christie took a bold step regarding this question, as it introduces a painting that claimed the art was created solely by AI without human intervention at all.[22] This action creates an uproar in the art world both among traditional artists and the group of artists working 'with' AI (AI artist hereafter) [23]. Traditional artists believe that the machine does not have intrinsic value; it lacks experience, emotion, and cultural understanding. Thus, if AI is allowed to be an author of the work, it might deteriorate the value of the art market.[24] On the other hand, AI artists who use AI as a vital 'tool' or 'assistant' of their work, agree with the traditional artists that artwork cannot be created solely by AI since a certain degree of intervention would be extremely necessary.[25] [26] Thus, the critical question is not about whether the AI can be the author, but whether the artist using AI could be considered artists as the tool lacks originality.

AI art generators are often criticized for their development process; as discussed above, generative AI art like Dell-E and Midjourney lacks originality since its output is close to the original artists; sometimes, the service providers even provide the users to target search with an artist name. Scoping down the artist's style by name plummeted the originality of the generative AI art further.[27] Moreover, the lack of originality in AI art resulted from its architecture, especially Generative Adversarial Networks (GAN) a model that can generate art by minimax game between generative and discriminate networks[28] [29], Creative Adversarial Networks (CAN) which is a special type of GAN that allows the AI to creative art by itself without or minimal human intervention [30]. While these two types of artificial Intelligence are almost entirely automated, there is still room for human artists' creativity to play;

Observers Can Discern Intentionality and Structure in Abstract Expressionist Art' (2015) 137 Cognition 154 <https://doi.org/10.1016/j.cognition.2014.12.009> accessed 10 September 2023.

[19] Annie Tubadji, Haoran Huang, and Don J Webber, 'Cultural proximity bias in AI-acceptability: The importance of being human' (2021) 173 Technological Forecasting and Social Change 121100 < https://doi.org/10.1016/ j.techfore. 2021.121100>

[20] Sofian Audry and Jon Ippolito, 'Can Artificial Intelligence Make Art without Artists? Ask the Viewer' (2019) 8(1) Arts 35<https://doi.org/10.3390/arts8010035>

[21] Marian Mazzone and Ahmed Elgammal, 'Art, Creativity, and the Potential of Artificial Intelligence' (2019) 8 Arts 26 <https://www.mdpi.com/2076-0752/8/1/26> accessed 1 November 2024.

[22] ibid 12.

[23] Kieran Browne, 'Who (or What) Is an AI Artist?' (2022) 55 (2) Leonardo 130 <https://doi.org/10.1162/leon_a_ 02092>

[24] ibid 8.

[25] ibid 12.

[26] ibid 23.

[27] ibid 12.

[28] Chris Donahue, Julian McAuley, and Miller Puckette, 'Adversarial Audio Synthesis' (the 7th ICLR, New Orleans, LA, USA, May 2019)

[29] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, *et al.* 'Generative adversarial nets' (2014) 27 Advances in Neural Information Processing Systems 2672 <https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf>

[30] Ahmed Elgammal, Bingchen Liu, and Mohamed Elhoseiny, 'CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms' (2017) <https://doi.org/10.48550/arXiv.1706.07068>

for example, they can choose the training data or teach how AI interprets texts or commands.[31]

As discussed above, AI art generators are often designed to be able to create creative work without human intervention at all. Thus, everyone can make the art without any expertise in the field. However, it is quite misleading to call this user an artist, but the AI artists are the experts who work with AI, a tool severely lacking originality, and could make it revolutionary by using their creativity to design and train AI.

III. CHALLENGES OF AUTHORSHIP AND ORIGINALITY

A fundamental aspect of copyright law is originality, a concept that has been challenging to define statutorily but remains crucial. This is because only works exhibiting a basic degree of this characteristic are eligible for protection.[32] None of the significant international copyright treaties provide a clear definition of originality or specify the degree of originality needed to obtain copyright protection.[33] Article 2(1) of the 'Berne Convention' focuses on 'protected works' and offers a representative, though not all-inclusive, list of works covered under this wide definition. The concept of 'original' is addressed in Article 2(3) and also in Article 14-bis, which pertains to cinematographic works. Additionally, Article 2 (5) introduces a pertinent aspect of originality, stating that works should be 'intellectual creations'. This raises the question of whether an 'intellectual creation' implies something produced by a human.

In this regard, many scholars argue that copyright should focus on safeguarding the 'results of human authorship' rather than being influenced by commercial factors.[34] This view is held because of worries that the essence of copyright might be overshadowed, diminishing the humanistic spirit of the 'Berne Convention'.[35] Ricketson asserts that acknowledging authorship is a basic human right for the creator of a work.[36] Limiting the notion of authorship to humans not only reinforces essential human values but also serves as a 'welcome reminder of human individuality and distinctiveness'.[37]

Nevertheless, some scholars argue that granting copyright protection to AI does not entirely eliminate its humanistic element but rather widens the gap between the created work and the author who developed the algorithm enabling the work (the programmer), effectively shifting the author's position in the creative process of the work.[38] Thus, the humanistic essence of the 'Berne Convention' continues to be protected, albeit in a less direct manner.[39]

Under the existing copyright protection systems, no country imposes an explicit prohibition on granting copyright to works created by artificial intelligence. However, it appears that many countries' laws do not permit non-human

[31] Zachary C. Lipton, 'The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery' (2018) 16(3) Queue 31 <https://doi.org/10.1145/3236386.3241340>

[32] Sam Ricketson and Jane C. Ginsburg, *International Copyright and Neighbouring Rights: The Berne Convention and Beyond* (Oxford University Press 2022); Lionel Bently et al., *Intellectual Property Law* (Oxford University Press 2022); W. R. Cornish, David Llewelyn, and Tanya Frances Aplin, *Intellectual Property: Patents, Copyrights, Trade Marks and Allied Rights* (Sweet & Maxwell 2023).

[33] Thomas Margoni, 'The Harmonisation of EU Copyright Law: The Originality Standard' (2016) SSRN Electronic Journal <https://doi.org/10.2139/ssrn.2802327>

[34] Jane C. Ginsburg, 'People Not Machines: Authorship and What It Means in the Berne Convention' (2018) 49(2) IIC - International Review of Intellectual Property and Competition Law 131; Sam Ricketson, 'The 1992 Horace S. Manges Lecture - People or Machines: The Bern Convention and the Changing Concept of Authorship' (1991).16(1) Columbia-VLA Journal of Law & the Arts 1

[35] Jane C. Ginsburg, 'People not Machines: Authorship and What it Means in the Berne Convention' (2018) 49 IIC 131 <https://doi.org/10.1007/s40319-018-0670-x>

[36] A. Michel, 'AI-Generated Creations: Challenging the Traditional Concept of Copyright, A Research into the Question of Works that are created by an Artificial Intelligence Program Have Copyright Protection in the Netherlands and the European Union' (Tilburg University 2018).

[37] ibid.

[38] Jesus Manuel Zatarain, 'The Role of Automated Technology in the Creation of Copyright Works: The Challenges of Artificial Intelligence' (2017) 31(1) International Review of Law, Computers & Technology 91

[39] Michel (n 36).

copyright. Majority of the scholars view that machines do not possess autonomy, and as a result, they are incapable of independently exercising their ownership rights based on their own free will.[40] Therefore, AI machinery is unable to assert its rights by initiating infringement lawsuits in a legal court.[41]

In the United States, for instance, the concept of originality requires human authorship.[42] This position is based on legal precedent, which establishes that copyright law protects only works that are the result of intellectual labor based on the creative abilities of the mind.[43] Moreover, the U.S. Copyright Office maintains a policy that they will not grant registration for works that are created solely by a machine or mechanical process that functions randomly or automatically, without any creative input or involvement from a human author.[44] Likewise, a recent case in Australia ruled that a work produced with the involvement of a computer cannot be protected by copyright because it was not created by a human.[45] In Europe, the Court of Justice of the European Union has also determined, particularly in its influential *Infopaq International v. Danske Dagblades Forening decision*[46] and *Football Dataco v Yahoo! UK decision*[47], that copyright only applies to original works that reflect the author's own intellectual creation. This is commonly interpreted to mean that an original work must reflect the author's individuality, which clearly implies that a human author is necessary for a work to be eligible for copyright protection. Under this view, a simple sketch of a stick figure drawn by a human hand holds more value in terms of copyright protection than the Next Rembrandt,[48] the artwork created by AI.[49]

If AI cannot be regarded as an author, it prompts the question of whether the ownership or creation of the AI should entitle the individual who developed it to be recognized as the author. It is undeniable that significant investment is needed to create AI, and ensuring that the creator has the chance to recover this investment is crucial. In reality, the majority of computer-generated artworks are significantly influenced by the creativity of the programmer. Therefore, a small number of countries, including Hong Kong (SAR), India, Ireland, New Zealand, and the UK, attribute authorship to the programmer.[50] In the UK High Court's ruling in *Nova Productions Ltd v Mazooma Games Ltd (Nova Productions),* the court assigned the rights to the programmer, as he was accountable for coding the program that allowed the computer to generate the different visuals in the frames that were under copyright protection.[51]In addition, the UK's copyright law, specifically section 9(3) of the Copyright, Designs and Patents Act (CDPA), embodies this approach. The CDPA stipulates that for a literary, dramatic, musical, or artistic work that is computer-generated, the UK law specifies that the person who made the necessary arrangements for creating the work is considered the author.[52] Dickenson suggests that determining who made the essential arrangements should involve identifying the individual who applied their skill,

[40] Petar Hristov Manolakev, 'Works Generated by AI – How Artificial Intelligence Challenges Our Perceptions of Authorship' (Master's Thesis, Tilburg University 2017) 38

[41] Kanchana Kariyawasam, 'Artificial Intelligence and Challenges for Copyright Law' (2020) 28(4) International Journal of Law and Information Technology 279

[42] Daniel J. Gervais, 'The Machine as Author' (2019) 105 IOWA L. REV. 2053

[43] *Feist Publications v Rural Telephone Service Company, Inc.* 499 U.S. 340 (1991)

[44] Maria Strong, 'Comments of the United States Copyright Office to the World Intellectual Property Organization: Impact of Artificial Intelligence on IP Policy: Call for Comments' *U.S. Copyright Office* (Washington, DC, 14 February 2020) <www.wipo.int/export/sites/www/about-ip/en/artificial_intelligence/call_for_comments/pdf/ms_usa_usco.pdf>

[45] *Acohs Pty Ltd v Ucorp Pty Ltd* [2012] FCAFC 16.

[46]*Infopaq International v. Danske Dagblades Forening* [2009] ECR 16 (Case C-5/08)

[47] *Football Dataco Ltd and others v. Yahoo! UK Ltd* and others [2012] EWCA Civ 1696 (Case C-604/10)

[48] In 2016, a collection of museums and scholars in the Netherlands revealed a painting named The Next Rembrandt, which was a novel artwork produced by an AI that had analyzed numerous pieces by Rembrandt Harmenszoon van Rijn, a Dutch artist from the 17th century.

[49] Daryl Lim, 'AI & IP Innovation & Creativity in an Age of Accelerated Change' (2018) 52 *AKRON L. REV.* 813

[50] Andres Guadamuz, 'Artificial Intelligence and Copyright' (*WIPO Magazine*, 2017) <www.wipo.int/wipo_ magazine/en /2017/05/article_0003.html>

[51] *Nova Productions Ltd v Mazooma Games Ltd & Ors* [2007] EWCA Civ 219, [2007] EMLR 427, [2007] RPC 25, [2007] BusLR 1032, [2007] 30(5) IPD 30032

[52] Copyright, Designs and Patents Act 1988, s 178.

labor, and judgment in making those arrangements.[53] It seems logical to consider that the term 'necessary arrangements' might pertain to the initial efforts involved in creating and developing the program.[54]

However, there may be scenarios where distinguishing between works created by humans and those generated by computers without human involvement becomes difficult. For instance, contemporary AI might possess the ability or autonomy to make independent choices, leading to minimal human input in the AI tool's results. In essence, despite a programmer typically being responsible for any algorithm used in AI, there are modern computers programmed without strict guidelines, enabling the machines to self-program.[55] When the role of the programmer diminishes, one might argue that the term 'arrangements' mentioned in section 9(3) of the CDPA encompasses both the act of programming and the investments made for it. Therefore, in cases of AI works independent of the programmer's involvement, the rights holder is deemed to be the entity that has made the financial investment.[56] This means that the term 'the person by whom the arrangements necessary for the creation of the work are undertaken' should be broad enough to include anyone intending to create AI-generated work.[57] This approach would initially focus on the main individual, the programmer, and in cases where the AI-generated work is fully autonomous, extend beyond the programmer's expertise to include the person who conceived the idea of creating such work.[58] In other words, when AI is entirely autonomous and able to produce work surpassing the programmer's understanding, some scholars suggest that for AI creations detached from the programmer's contribution, the rights may belong to the entity responsible for the financial investment.[59]

It is important to recognize that with more advanced AI programs capable of learning and adapting independently, this interpretation of 'necessary arrangements' might be pushing the concept in a broad manner and it is likely that UK courts would interpret the provision in a broad manner to fulfill its intended purpose.[60] It should be noted, however, that, given that AIs employ deep-learning algorithms, the gap between the creator who implemented the initial 'necessary arrangements' and the ultimate output becomes even greater.[61]

Furthermore, the CDPA defines a computer-generated work as one that is generated by a computer in circumstances where there is no human author involved.[62] The protection period for such works is 50 years from the date of creation.[63] The goal of this provision is to make an exception to the requirement for human authorship by acknowledging the effort involved in developing a program that is capable of producing works, even though the machine undertakes the creative process.[64] The way the UK law designates a human as the author of an AI-generated work separates authorship from creativity, which contradicts the modern approach to originality in copyright law where authorship and creativity are linked.[65] In addition, the Intellectual Property of the UK further clarifies that the concept of 'joint authorship' does not apply to works co-created by humans and AI systems because computer-generated works have

[53] Julia Dickenson, Alex Morgan, and Birgit Clark, 'Creative Machines: Ownership of Copyright in Content Created by Artificial Intelligence Applications' (2017) 39(8) *E.I.P.R.* 457
[54] Michel (n 36).
[55] Elsa Sayagh, 'Can Artificial Intelligence Be More Creative than Humans?' (*Welcome to the Jungle*, 02 April 2019) <www.welcometothejungle.com/en/articles/en-can-artificial-intelligence-be-more-creative-than-humans> accessed 27 January 2024.
[56] Madeleine De Cock Bunning, 'Buning Artificial Intelligence and The Creative Industry: New Challenges For The EU Paradigm For Art And Technology By Autonomous Creation' in *Research Handbook on The Law of Artificial Intelligence* (Edward Elgar Publishing 2018) 511–35
[57] Kariyawasam (n 41).
[58] ibid.
[59] ibid.
[60] Dickenson, Morgan, and Clark, (n 53).
[61] Michel (n 36).
[62] ibid.
[63] CDPA, s 12(7).
[64] Intellectual Property Office, 'Artificial Intelligence Call for Views: Copyright and Related Rights' (*GOV.UK*, 23 March 2021)https://www.gov.uk/government/consultations/artificial-intelligence-and-intellectual-property-call-for-views/artificial-intelligence-call-for-views-copyright-and-related-rights> accessed 26 February 2023.
[65] ibid.

no human author.[66] This creates ambiguity regarding the status of AI-assisted works, so there is a need for clarification of these provisions. While it was anticipated that other nations would also offer comparable legal protection for works created by artificial intelligence, at present, only a few countries, aside from the UK, take the same approach.

Another different viewpoint is that AI could be the holder of copyright for creative works produced solely by AI, considering AI as a corporate entity. According to Shawn Bayern's suggestion, assigning a computer system to manage a limited liability corporation (LLC) may give artificial intelligence the status of a legal person capable of owning property.[67] Acevedo argues that corporations have organizational structures that include human directors who are empowered to make decisions on behalf of the organization.[68] In contrast, artificial intelligence does not have these conventional features. However, it could be argued that the individual who programmed the AI acts as the director, and the AI itself functions as the Chief Executive Officer (CEO).[69] In doing so, Acevedo suggests the U.S. copyright office apply the structure of the work made for hire concept or to treat AI as a joint author with programmer according to the joint work concept.[70]

In 2019, the International Association for the Protection of Intellectual Property (AIPPI) conducted a survey among its members to determine if works generated by artificial intelligence should be protected. The responses received revealed a range of opinions on the matter that, while the UK Group proposed a new 25-year right to protect AI-generated works and acknowledge the investment of AI developers, other respondents believed that copyright protection should be reserved only for works that result from human creativity.[71] The resulting AIPPI Resolution, however, still emphasizes the importance of human intervention and originality in the copyright protection of works.[72]

IV. INCENTIVES FOR CREATION IN THE DIGITAL AGE

Originality, although an important concept when considering granting copyright protection to AI-generated works, does not serve the purpose to organize the difficult debate on the topic. Aiming to derive a solution by using the concept of originality leads to further discussion on whether AI art itself is a form of art, or as put by Coeckelbergh, 'can machines create art?'[73] Studies that frame the issue in this way usually attempt to consider the definition of art itself then relate back to theoretical concepts of copyright.[74] The key argument usually asserts that AI is not creative, or more specifically, not creative in the way humans are[75], thus not worthy of copyright protection.[76]

Meanwhile, another side of the ongoing debates focus on the shift of human contribution and its value. This usually concerns the changing degree, more precisely the reducing effort and control of humans[77], perhaps a reflection of determination to

[66] ibid.
[67] Roman V. Yampolskiy, 'Could an artificial intelligence be considered a person under the law?' (*The Conversation*, 5 October 2018) <https://theconversation.com/could- an-artificial-intelligence-beconsidered-a-person-under-the-law-102865>
[68] Veronica Acevedo, 'Original Works of "Authorship": Artificial Intelligence as Authors of Copyright' (Student Works, Seton Hall University 2022)
[69] ibid.
[70] ibid.
[71] Intellectual Property Office (n 64).
[72] Jan Bernd Nordemann, 'AIPPI: No Copyright Protection for AI Works without Human Input, but Related Rights Remain' (*Kluwer Copyright Blog*, 21 November 2019) <https://copyrightblog.kluweriplaw.com/2019/11/21/aippi-no-copyright-protection-for-ai-works-without-human-input-but-related-rights-remain/>
[73] Mark Coeckelbergh, 'Can Machines Create Art?' (2017) 30 Philosophy & Technology 285 <http://link.springer.com/10.1007/s13347-016-0231-5> accessed 7 November 2024.
[74] Jiang and others (n 8).
[75] Ujué Agudo and others, 'Assessing Emotion and Sensitivity of AI Artwork' (2022) 13 Frontiers in Psychology 879088 <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.879088/full> accessed 7 November 2024.
[76] Samuel Scholz, 'A Siri-Ous Societal Issue: Should Autonomous Artificial Intelligence Receive Patent or Copyright Protection?' (2020) 11 Cybaris® <https://open.mitchellhamline.edu/cybaris/vol11/iss1/3>.
[77] Uwe Messer, 'Co-Creating Art with Generative Artificial Intelligence: Implications for Artworks and Artists' (2024) 2 Computers in Human Behavior: Artificial Humans 100056

carefully create works as what would be imagined for other forms of works. Such concerns are linked with anxiousness that generative AI devalues the value of labor and skills of human artists[78], while the more hopeful side proposes that AI-generated arts may in fact induce the public to value the works of human artists more, if only they know which are purely created by humans and which not.[79]

Above are examples of how recent scholarship rightly raises several concerns, yet still struggles to arrive at solutions as the topic is framed as a 'problem' where generative AI disrupts and changes the creative world as we know. However, it is important to note that creativity and the act of creation itself are not the same thing. This section therefore attempts an analysis that acknowledges generative-AI-involved creations are not to be viewed merely at the level of the resulting work, but rather at the level of creative process.[80] Here, therefore, the goal is to extract the roles of several actors who have their ways and degrees of contribution, and carefully restart from the starting point of incentive to create[81]

The complexity of creation with generative AI is not only that machines are an integral part of the creative process, but also that several human actors are involved, each with some level of contribution that may be interpreted as putting effort or deserving rewards. Creation with generative AI involves contribution from several author-like actors, including (1) the programmer (the one who created the generative AI), (2) the investor[82] (the person or entity that supports the operation through funding and/or infrastructure such as server), (3) the AI system, (4) the curator[83] (the human who trains the generative AI to create the most desired or marketable style of work), and (5) the creator (the human who inputs the command for the AI to generate). In some cases, the human(s) involved can fulfill multiple roles in a creative process. For example, a programmer can also be a curator.

Discussions on legal options for AI-involved creations are surrounded by making a choice to not grant authorship at all or giving authorship to some actor involved in the creative process. The main obstacle of analysis is leaving the root of criteria for copyright protection unclarified, i.e. as mentioned above, actors in the creative process show some level of contribution. In the digital age, the value chain of creation no longer allows a simple method of identifying a single actor with notable contribution, as it is 'it is often impossible to identify the particular (human) authors responsible for expressing or compiling part of the work.' [84] To what degree must an actor in generative AI process contribute, and when exactly does an actor deserve or *require* authorship?

At the same time, incentive for creation is commonly used as justification and basis of copyright laws.[85, 86] Conventionally, it is viewed that copyright works through granting monopoly over the creator's work[87], and authorship is considered to be the source of creative incentive[88]

<https://linkinghub.elsevier.com/retrieve/pii/S2949882124000161> accessed 15 November 2024.

[78] Jiang and others (n 8).

[79] C Blaine Horton Jr, Michael W White and Sheena S Iyengar, 'Bias against AI Art Can Enhance Perceptions of Human Creativity' (2023) 13 Scientific Reports 19001 <https://www.nature.com/articles/s41598-023-45202-3> accessed 15 November 2024.

[80] Mazzone and Elgammal (n 21).

[81] Kateryna Militsyna, 'Human Creative Contribution to AI-Based Output – One Just Can('t) Get Enough' (2023) 72 GRUR International 939 <https://academic.oup.com/grurint/article/72/10/939/7241907> accessed 3 January 2024.

[82] Kariyawasam (n 41).

[83] Ramya Srinivasan and Kanji Uchino, 'Biases in Generative Art: A Causal Look from the Lens of Art History' (FAccT '21: 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual Event, Canada, 2021) 41–51 <https://doi.org/10.1145/3442188.3445869>

[84] Jani McCutcheon (n 2) 915–69.

[85] Sara K Stadler, 'Incentive and Expectation in Copyright' (2007) 53 (3) Hastings Law Journal 433

[86] Shyamkrishna Balganesh, 'Foreseeability and Copyright Incentives' (2009) 122 (6) Harvard Law Review 1569

[87] Seana Valentine Shiffrin, 'The Incentives Argument for Intellectual Property Protection' Axel Gosseries, Alain Marciano, and Alain Strowel Zin (ed) *Intellectual Property and Theories of Justice* (London: Palgrave Macmillan, 2008) <https://doi.org/10.1057/978-0-230-58239-2_5>

[88] James Campbell, 'Authorship, Incentives for Creation, and Copyright in the Digital 21st Century' (2007) 43(1) Proc. Am. Soc. Info. Sci. Tech. 1 <https://doi.org/ 10.1002/ meet.1450430168>; Tim Worstall, 'Copyright Is About Incentives to Innovation, Not Justice: What Incentive Does Naruto Need?' (Forbes, 07 January 2016) <www.forbes.com/sites/timworstall/2016/01/07/copyright-is-about-incentives-

However, such views are arguably rooted and constrained within only some philosophical bases, namely utilitarian and Lockean theories. In economic perspectives, incentives are necessary for inventions and creative works, and there are other types on incentives than intellectual property protection that one might put contribution into some creation[89]. Hence, in this analysis, the scope of incentives must be broadened, from incentives only granted from copyright itself to other incentive schemes, to truly examine why each actor creates. To derive the fitting solution for actors in creative process with generative AI, it is necessary to revisit philosophical and economic justifications of copyright allocation, and apply them to the actors involved. Fisher[90] and Former[91] provide useful overview and comparisons of the four philosophical grounds for copyright, and intellectual property regimes in general. The four guidelines can be first classified into utilitarian basis and moral rights (non-utilitarian) basis[92], where the moral rights basis is further distinguished into labor and natural rights approach, personality theory, and social planning theory. [93, 94] Here, each philosophical approach is not discussed for their validity or issues, but rather utilized for application to clarify incentives required for creative process.

(a) The utilitarian basis of incentive for creation is famously embodied in the US Constitution, which empowers the Congress '[t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.'[95] Intellectual property protection is justified by maximization of net social welfare[96], where providing incentives to create and innovate would generate artistic and technological progress that benefits the society.

(b) The labor and natural rights approach, prominently known as Lockean theory of intellectual property[97], is derived mainly from the works of Nozick and Locke.[98] Scholars derive grounds for intellectual property protection based on the Lockean idea of property.[99] The approach draws upon Locke's concept of *proviso*[100], which stipulates that any property in the commons, once mixed with labor of a person, that person is legitimized to acquire the property as their own. When applied to copyright, efforts and resources of the creator are compared with labor, concluding that the creator is entitled with the natural rights in their property, the works they created.[101]

(c) The personality theory[102], also known as personhood theory, is derived loosely from the works of Kant and Hegel, which focuses on the moral value of human's 'will' and personhood expressed through creations – that private property rights are crucial for fulfillment of fundamental human needs to express their personality.[103] In this view, 'rights exist to effectuate a person's basic human desires'.[104] This leads to the argument that the creators morally deserve rights in their work, considering the

to-innovation-not-justice-what-incentive-does-naruto-need/ > accessed 19 March 2023.

89 Nancy Gallini and Suzanne Scotchmer, 'Intellectual Property: When Is It the Best Incentive System?' (2002) 2 Innovation Policy and the Economy 51 <https://doi.org/10.1086/653754>

90 William Fisher, 'Theories of Intellectual Property' (1987) <https://cyber.harvard.edu/people/tfisher/iptheory.pdf>

91 Jeanne C. Fromer, 'Expressive Incentives in Intellectual Property (2012) 98(8) Virginia Law Review 1745

92 ibid.

93 Fisher (n 90).

94 Justin Hughes, 'The Philosophy of Intellectual Property' (1988) 77 Geo. L.J. 287 <https://cyber.harvard. edu/ IPCoop/ 88hugh.html>

95 U.S. CONST., art. I, § 8, cl. 8.

96 Fisher (n 90).

97 Adam D. Moore, 'A Lockean Theory of Intellectual Property Revisited' (2012) *SSRN Electronic Journal* <https://doi.org/ 10.2139/ssrn.2099073>

98 Fisher (n 90).

99 Mala Chatterjee, 'Lockean Copyright versus Lockean Property' (2020) 12 Journal of Legal Analysis 136 <https://doi.org/10.1093/jla/laaa002>; Adam D. Moore, 'A Lockean Theory of Intellectual Property Revisited' (2012) 50 San Diego Law Review <https://ssrn.com/abstract= 2099073>

100 John Locke, 'Second Treatise of Government' in Mitchell Cohen (ed) *Princeton Readings in Political Thought: Essential Texts since Plato - Revised and Expanded Edition* (Princeton University Press 2018) <https://doi.org/ 10.2307/j.ctv19fvzzk>

101 Fisher (n 90).

102 Hughes (n 94).

103 Fisher (n 90).

104 Hughes (n 94).

virtue of their works and how their creation reflect their experience, knowledge, personhood[105], and may be extended to skills they have learned and assembled overtime. Intellectual property regime, therefore, exists to 'shield the creators from appropriation or modification of artifacts through which authors and artists have expressed their "wills"', or 'to create social and economic conditions *conducive* to creative intellectual activity.'[106] Some scholars point out overlap between Lockean and personality theories, that application of one's labor is a form of expression.[107] Here, these shall be distinguished, that the personality theory points towards judgment and decisions of the creator, while Lockean property points towards applying skills and labor.

(d) Social planning theory posits that artistic creations and protecting them with property rights are linked to shaping the society, by promoting 'the achievement of a just and attractive culture.'[108] Copyright is linked with preservation and enrichment of artistic tradition, to create 'just and attractive culture'.[109] This theory, although different in the details of functioning of copyright, shares some common grounds with the utilitarian basis.

With the four justifications of copyright clarified above, it is now possible to further examine the nature of incentives as would be viewed by scholars of each theoretical approach. Here, another dimension that allows clearer insight is viewing how incentives drive creators – through creating motivations. They can be categorized into 'intrinsic' and 'extrinsic' motivations.[110, 111, 112]

Intrinsic motivations are formed internally, thus we observe examples of creators, for example, as raised by Moglen[113],Mozart composed *The Magic Flute* knowingly he would not be paid. Extrinsic motivations, on the other hand, depend on what system offers to the creators. With these laid out, factors that make creators create can now be clarified for each theoretical approach.

The utilitarian basis provides the clearest and the most prevailing concept of source of incentive – copyright as 'reward'[114], which directly performs as an incentive, in line with economic theories of incentive. The motivation in the utilitarian approach is completely extrinsic. As concisely put by Snow, it is the 'State's intervention to incentivize creativity with subsequent income from excluded right to the works.'[115] The creators can form expectations of their rewards, through exclusive rights on the works and exploitation.[116]

The labor theory, or Lockean property approach, would stipulate that ensured fairness motivates the creators, as they are 'ensured the entitled level of fairness for their labor'[117] through recognition of authorship. There is a slight overlap with utilitarian basis here, as such level of fairness is linked with possible reputation and income as entitled acknowledgement of hard work. This makes motivation based on the Lockean approach a combination of mainly intrinsic, and partially extrinsic motivations.

The remaining two approaches based on moral rights, personality theory and social planning theory involve purely intrinsic motivations. Personality theory implies the fulfillment of will, the fundamental human desires. The motivations are linked to personhood, including self-realization as a social being, self-realization as

---

[105] Fromer (n 91).
[106] Fisher (n 90).
[107] Hughes (n 94).
[108] Fisher (n 90).
[109] ibid.
[110] Ece Gurler, 'Assessing the Role of Motivational Factors in Facilitating Artists' Personal and Professional Development' (2021) Critical and Creative Thinking Capstones Collection 393 <https://scholarworks.umb.edu/cct_capstone/393>
[111] David M. Kreps, 'Intrinsic Motivation and Extrinsic Incentives' (1997) 87(2) The American Economic Review 359
[112] Gary Charness and Daniela Grieco, 'Creativity and Financial Incentives' (2019) 17(2) Journal of the European Economic Association 454 < https://doi.org/10. 1093/jeea/ jvx055>
[113] Eben Moglen, 'Anarchism Triumphant: Free Software and the Death of Copyright' (1999) 4(8) First Monday <https://doi.org/10.5210/fm.v4i8.684>
[114] Ioannis Lianos, 'A Regulatory Theory of IP: Implications for Competition Law' (2008) CLES Working Paper Series 1/2008<https://www.ucl.ac.uk/cles/sites/cles/files/cles-1-2008new.pdf>
[115] Ned Snow, 'Moral Bars to Intellectual Property: Theory & Apologetics' (2021) 28(1) UCLA Entertainment Law Review <https://doi.org/10.5070/LR828153857>
[116] Stadler (n 85).
[117] Snow (n 115).

an individual, and security and leisure.[118] Social planning theory would posit that the incentive lies in contributing to the society by enriching artistic tradition, and contributing to distributive justice, where creators do their part to contribute to the society, and the society has a just opportunity to enjoy the works.[119]

Beyond motivations derived from incentives based on the four justifications of copyright, three developments in how incentives are understood must be mentioned. Market factors such as prizes, subsidies, and regulation[120], and changing market trends provide incentive schemes other than copyright. Further, 'negative space scholarship' of intellectual property has been proposed in recent years, challenging ways of grasping the necessity of copyright protection as incentive to create.[121] For example, low-cost innovation may not require copyright protection or authorship attribution for the creators to maintain motivations and for the society to benefit, and creativity in the market may be accelerated through creativity-enhancing copying, such as fast fashion industry where it is common for makers to copy each other.[122] Third, specific to creations involving generative AI, creative process increasingly involves lower cost of creation, higher volume of work production, and AI-assisted ideation.[123, 124]

[118] Jeremy Waldron, *The Right to Private Property* (Oxford 1990; online edn, Oxford Academic 2011) <https://doi.org/10.1093/acprof:oso/9780198239376.001.0001>

[119] Fisher (n 90).

[120] Peter S. Menell, 'Intellectual Property: General Theories' (1999) 2 Encyclopedia of law and economics 129 <https://core.ac.uk/reader/7280110>

[121] Christopher Jon Sprigman, 'Copyright and Creative Incentives: What Do(n't) We Know? in Rochelle Cooper Dreyfuss and Elizabeth Siew-Kuan Ng (ed), *Framing Intellectual Property Law in the 21st Century* (1st ed, Cambridge University Press 2018) 32–61 <https://doi.org/10.1017/9781316471647.003>

[122] ibid.

[123] Anne Ploin, Rebecca Eynon, and Isis Hjorth *et al.* 'AI and the Arts: How Machine Learning Is Changing Artistic Work' (Report from the Creative Algorithmic Intelligence Research Project, University of Oxford, UK: Oxford Internet Institute, 2022)<https://www.oii.ox.ac.uk/wp-content/uploads/2022/03/040222-AI-and-the-Arts_FINAL.pdf>

[124] Anjan Chatterjee, 'Art in an Age of Artificial Intelligence' (2022) 13 Frontiers in Psychology 1024449<https://doi.org/10.3389/fpsyg.2022.1024449>

Authorship must be allocated to the actor that requires incentives to create, and without whom the creation process would not occur, i.e. actors whose role is indispensable for the creation process. As summarized in Table 1, actors are listed, and the necessity of copyright as incentives for each actor is proposed by clearly distinguishing the four theoretical frameworks: utilitarian, labor, personality, and social planning. For Lockean (labor) framework, skill and labor are considered as the key representations of the concept, while for personality framework, judgment of the actor is considered as key representation of expression of personal will. It is found that programmers are most likely justified to require copyright by the theoretical frameworks, while the AI systems remain the least likely to be justified, as to be further elaborated below.

4.1 Programmer

The developers of artificial intelligence, require incentive for creation through copyright based on all theoretical frameworks. Programmers put significant effort into creating the working systems, for which they invest their time and workload to produce generative algorithms. They may be considered the 'mastermind'[125] with creative input[126] in the form of programming – without whom no process of generative AI would ever take place. The importance of their role in computer-created work may even be extended that they are the only individual in the creative process who contributes enough intellectual effort to satisfy justification for intellectual property claim.[127]

On a utilitarian basis, a programmer is the key person indispensable for advancement of the creative work market, as noted in the UK court case of *Nova Productions Ltd v Mazooma Games*

[125] Jane C. Ginsburg and Luke Ali Budiardjo, 'Authors and Machines' (2018) SSRN Electronic Journal <https://doi.org/10.2139/ ssrn.3233885>

[126] Andres Guadamuz, 'Artificial Intelligence and Copyright' (WIPO, October 2017) <www.wipo.int/wipo_magazine/en/2017/05/article_0003.html>

[127] Evan H. Farr, 'Copyrightability of Computer-Created Works' (1989) 15(1) Rutgers Computer and Technology Law Journal 63

Ltd as 'the person by whom the *arrangements necessary for the creation* of the works were undertaken'. [128]. The programmer requires rewards as incentives, which can be provided in form of income – for which the increasing income from copyright protection would then incentivize the programmer to create more for the progress of the society.

With the Lockean framework, developers put in labor and produce progress with a key element of generative AI works - the algorithm. In the US, the Congressional Research Service published an article that recognizes the effort and creative involvement of the programmer is not dismissible, to the extent that the developers might have 'a stronger claim to some form of authorship than the manufacturer of a camera'.[129] Their virtue of works, how their skills and labor create works with high impact, i.e. innovating AI systems that can enable further creations[130] can be acknowledged by copyright.

Meanwhile, two other frameworks, personality and social planning frameworks also support necessity of creative incentives from copyright for programmers. By exercising their significant amount of creative involvement and decisions, the programmers fit into the personality framework of self-expression and self-realization through the systems they develop. In social planning perspective, they must be incentivized to create systems that can enable further creations[131].

4.2 Investors

An investor initiates and/or supports resources for development of generative AI systems. The investors are usually represented by companies or organizations that fund the programmers or operate infrastructure, such as servers and data centers for processing of generative AI systems. If an investor is viewed as the employer of programmers and 'owner' of a generative AI system, an analogy can be made with works made for hire. In the United States, when an artist is hired or commissioned some work, the employer is deemed the author of works made for hire, not the artist.[132] Although currently, in context of generative AI works, work-made-for-hire concept is used mainly as analogy for discussions regarding authorship of creators (end users)[133,134,135,136,137,] this article proposes it should be used as an analogy for investors, which is less considered in creative processes with generative AI systems.

On a utilitarian basis, similar to programmers, investors play a key role in 'making necessary arrangements,' even though they may not be 'masterminds' of the creations themselves. Investors likely seek financial return for investment and operations. This implies possible necessity of income as reward for investors to ensure the progress of the society. However, despite the necessity, allocating authorship to the investors must be considered carefully against allocating authorship to the programmer, as this reflects the works-made-for-hire-concept, which can be found to conflict with moral rights frameworks.[138]

---

[128] *Nova Productions Ltd* (n 51).
[129] Christopher T. Zirpoli, "Generative Artificial Intelligence and Copyright Law," (CRS Legal Sidebar, 22 February 2023) <https://crsreports.congress.gov/product/pdf/LSB/LSB10922>
[130] Ryan Benjamin Abbott, 'Artificial Intelligence and Intellectual Property: An Introduction' in Edward Elgar (ed), *Research Handbook On Intellectual Property And Artificial Intelligence,* (Ryan Abbott, ed., Forthcoming) (SSRN Electronic Journal, 2022) <https://doi.org/10.2139/ssrn.4065150>
[131] ibid.
[132] Rochelle Cooper Dreyfuss, 'In Praise of an Incentive-Based Theory of Intellectual Property Protection' in Rochelle Cooper Dreyfuss and Elizabeth Siew-Kuan Ng (ed), *Framing Intellectual Property Law in the 21st Century* (1st ed, Cambridge University Press, 2018) 22 <https://doi.org/10.1017/9781316471647.002>
[133] Giorgio Franceschelli and Mirco Musolesi, 'Copyright in Generative Deep Learning' (2022) 4 Data & Policy e17 <https://doi.org/10.1017/dap.2022.10>
[134] ibid.
[135] Gönenç Gürkaynak et al., 'Questions of Intellectual Property in the Artificial Intelligence Realm' (2017) 3(2) The Robotics Law Journal 9 <https://ssrn.com/abstract=3295747>
[136] Helene Margrethe Bøhler, 'EU Copyright Protection of Works Created by Artificial Intelligence Systems' (Master's thesis, The University of Bergen, 2017) <https://hdl.handle.net/1956/16479>
[137] Kalin Hristov, 'Artificial Intelligence and the Copyright Dilemma' (2017) 57(3) IDEA: The IP Law Review <https://ssrn.com/abstract=2976428>
[138] Dreyfuss (n 132).

With the Lockean property framework and social planning framework, the investors may, in some cases, require incentives from authorship. For example, they may spend business skills and working time to build up organizations and assemble funding for development. They may also aim to contribute to the cultural richness of the society by promoting development of systems that can then create more works. However, these scenarios for incentives of investors using the Lockean property framework and social planning framework themselves indicate that motivations of the investors can vary. In the perspective of personality framework, it is unlikely that the investors exercise creative judgment that affect the creative process, or make any creative decisions that can be considered as expressions of their will or personhood.

4.3 AI

As a machine and algorithm, an artificial intelligence system does not perceive nor require incentives given through copyright as financial rewards (utilitarian approach). It also does not require incentive to perform tasks with labor and skills (Lockean approach), as the system's processing and learning are all operated as part of non-perceiving machine, i.e. AI does not feel exhausted after generating or learning a series of works. In terms of social contribution, AI merely performs the tasks as instructed and does not possess the idea of enriching the society with works it generates (social planning theory). Finally, the machine has no will and personhood as human creators do, and do not endeavor towards self-realization through creation (personality theory), although in the future, if AI is to be somehow granted legal personality[139], and it can be proven that the personhood of the system must be incentivized for expression, it may become possible to argue for justification of copyright as incentive for AI systems.

4.4 Curator

Curators participate in the creative process of works with generative AI by choosing previous works, either with an objective to develop the basic learning of the model itself or to ensure the model is trained towards 'marketable' creations. On a utilitarian basis, they may require income as incentive through copyright, but it is not absolutely necessary, as they may already receive income in form of wage. They improve the quality and/or marketability of the generated works, but they are not essential to the process when compared to programmers, as the person by whom the *arrangements necessary for the creation* of the works were undertaken.[140] With the Lockean property framework, it is similarly uncertain if they necessarily require incentive through copyright to acknowledge their skills and labor. In some cases, the curator may commit a significant amount of time to train the system, yet in many cases they may just be pouring sample works into the system and see what the system spits out eventually. On the other hand, the personhood framework justifies the necessity, as they indeed exercise creative judgment. For example, a curator for a generative AI system for pictures may train the model with certain Renaissance pieces, so that the system can learn to generate Renaissance style works. Even in cases where they only curate works for marketability, they exercise creative judgement on what would be marketable. Finally, with social planning theory, it is difficult to determine justification for curators as their objectives are uncertain – to contribute to cultural richness or to satisfy the creators who plan to use the trained system for their works.

4.5 Creator

Creators are the end users of the generative AI systems, the ones at the end of the creative process where the end product is made. In pre-generative-AI creative processes, creators are viewed as the essential actor that under 'romantic' notions spend immense amounts of time and resources to create.[141] However, with generative AI, the role of creators have changed, and with a variety of their newly defined roles as well. Creators can now 'divide' creative tasks to the AI, ranging from ones who use the generative AI system as an assistance, to ones who type in

[139] Acevedo (n 68).

[140] *Nova Productions Ltd* (n 51).

[141] Campbell (n 88).

commands and take the generated work as the end product. It is difficult to identify which pieces are made with what extent of AI-assistance, and it is possible that AI will be given more roles in the future[142], effectively reducing the cruciality of creative contribution of the human creators. With the assistance, creators are able to produce a higher volume of work, at a faster pace than ever before.

On a utilitarian basis, the necessity of copyright to bestow incentives upon the creators through income becomes case-by-case basis. In contrary with romantic notions of creation[143], where their intense commitments are used as justification for income, they can now sell their works at lower prices, but with higher volume. This fits economically well for creators as business actors, where they have lower cost and more sales. Meanwhile, with unpredictability of results[144], it is arguable that they are not the ones who make necessary arrangements for creation. If one argues for incentives for their contribution by inserting appropriate commands amidst unpredictability, the incentive should be for expert users of generative AI tools to share their techniques may be beneficial for the society and progress of creative industry[145], rather than to further augment incentives to input commands.

However, as mentioned above, AI involvement can vary in each case, so there may still be cases where they are truly the ones who make necessary arrangements. For example, in a report on machine learning and artistic works, the cover of the report is drawn by a professional illustrator, who in one section of the report reveals that utilizing generative AI was helpful to gain inspiration and to create something new, as artists usually draw on same things and end up finding difficulty to create completely unique pieces.[146] In such a case, where the artist gains inspiration from AI, then further creates their own work, it is possible to still argue for the artist.

In Lockean perspective, it becomes extremely difficult to justify the necessity of authorship to acknowledge their skills and labor. Creators in generative AI processes inevitably make smaller efforts to create and to gain skills which enables production of their work. They likely spend lower financial cost, labor, and time to learn, ideate, and create.

For example, in the history of music, from Bach to Mozart and Haydn, then onto early works of Beethoven, each composer received influence from other admirable composers. All such processes happened in their human brains, for example, Mozart's Great Mass in C Major (K.427) contains Bach and Händel's influence resulting from Mozart's 'intensive studies of Baroque compositions.'[147] If Mozart were alive today, he might not need to study Bach's fugues to compose with inspiration from Bach's works, as he would soon be able to input into a generative AI, perhaps, 'a mass setting in C major lasting about an hour, with techniques of Bach'. After receiving the output, the AI-using Mozart can then choose to either develop the piece further himself or launch a premiere immediately. Although this example seems extreme, this can be now observed where some pieces such as 'Happy Birthday as Mozart would have written' are created and posted on social media by content creators.

It is difficult to determine the extent of AI involvement in such pieces, consequently also the effort and authorship. The artists can therefore forfeit their effort by stating that they were inspired by previous compositions, and with their human brains created the pieces, even though they may have indeed used generative AI in some or almost whole part of the process. Another

[142] Kariyawasam (n 41).
[143] Campbell (n 88).
[144] Gerald Spindler, 'AI and Copyright Law: The European Perspective' in Larry A. Di Matteo, Cristina Poncibò, and Michel Cannarsa (ed), *The Cambridge Handbook of Artificial Intelligence* (1st ed., Cambridge University Press, 2022) 257–70 <https://doi.org/10.1017/9781009072168.025>
[145] Ann Kristin Glenster, *Policy Brief: Generative AI* (Cambridge, UK: Minderoo Centre for Technology and Democracy 2023) <https://doi.org/10.17863/CAM.101918>
[146] Ploin, Eynon, and Hjorth *et al.* (n 123).
[147] Wolfgang Amadeus Mozart, "Preface," in Monika Holl and Karl-Heinz Köhler (ed), *Mozart Mass in C Minor (K.427 / K.417a)* (Baerenreiter Germany 2018) <https://www.barenreiter.co.uk/ prefaces/ 9790006 202232_ Innenansicht .pdf>

example would be generative AI systems that embed one's voice into a music cover, such as AI Voice, which by just learning the creator's voice, the AI can generate a cover version of any song with that voice, pitching all the correct notes. Having less cost by not having to intensively study the whole techniques of music is, in itself, incentive to create. It enables new pieces to be created more rapidly, and with more volume.

Negative space scholarship provides an example of the fast-fashion industry, which resembles the situation of generative AI arts, in terms of using massive amounts or rapidly created amounts of other works as inspiration. The mechanism of creativity-enhancing copying is items drive creativity for firms and designers to innovate new products that customers can enjoy at greater pace and volume, without requiring creative incentives from authorship and copyright protection.[148]

With a personhood framework, it is still possible to justify the necessity of copyright to provide incentives to create. The creators, even though with less effort, exercise their will with creative judgements to input a certain command. AI used in creative industries as co-creator or a tool help augment human their creativity[149], most likely with what they at least broadly creatively envision to make, with AI as source of inspiration of how to exactly create. Although some reject the concept that inspiration through learning process and experience of conventional artists cannot be considered the same as using generative AI as source of inspiration[150], it is in fact shortening the learning and experiencing steps, and sensible creators still must know what their own (or the society's) preference of 'good' creation is and how they should further develop their works. Similarly, another rejection is the concept that using generative AI as a source of inspiration would mean using only previous works as reference, which would lead humanity to 'an ouroboros where nothing new is truly created, a stale perpetuation of the past...'[151] It is again important to take notice that creators, just like Mozart, learn from their exposure to surroundings and previous creators. This process of 'contemporary generative art' can be considered as bricolage, 'where ideas are developed through playful experimentation with existing tools and techniques', and it is not unique to generative AI age – artists in history, such as Leonardo da Vinci, constantly experiment with ideas, with many not finished as final piece of work.[152]

It may be possible that as the technology and market of AI-generated works mature, society would perceive AI-generated works as a novel field of creative work. The situation would be similar to how photography gradually became perceived as artwork, where cameras are perceived as tools. However, beyond the conventional concept of originality, the key difference between generative AI and cameras at the current stage is 'free choice'[153] of the artists, in other words whether the artists can predict the output of their intended creation. If the day ever comes when creators indeed exercise free choice, in this context their will and personhood, justification with personhood framework would become even stronger. Meanwhile, similarly, in social planning perspective, it depends on the degree the creators can express their will to create, that it would be possible to identify their intentions to contribute to the cultural richness of the society.

V. CONCLUSION

The dilemma of balancing creative incentives with originality is central to the discourse on AI and copyright. On the one hand, copyright law has traditionally functioned as a mechanism to encourage creative endeavors by providing exclusive rights as incentives. This incentive structure is predicated on the notion that creativity needs nurturing through the promise of financial and reputational gain. However, the advent of AI in creative domains complicates this narrative.

[148] Sprigman (n 121) 46.
[149] Daniele and Song (n 12).
[150] Jiang and others (n 8).
[151] ibid.
[152] Dejan Grba, 'Forensics of a Molten Crystal: Challenges of Archiving and Representing Contemporary Generative Art' (2019) 8 ISSUE Annual Art Journal 3, <https://doi.org/10.33671/ISS08GRB>
[153] Spindler (n 144) 261.

AI's ability to produce works that challenge our notions of originality calls into question the very foundations of copyright law. AI-generated works, while novel and capable of evoking aesthetic responses, often stem from algorithms processing vast datasets of existing human-created content. This raises an important question whether these works can be considered original in the legal sense. Furthermore, if the primary objective of copyright is to incentivize human creativity, where does that leave AI-generated works that do not require the traditional forms of human labor and ingenuity?

A pivotal aspect of adapting copyright law to AI involves redefining legal definitions and examining precedents. Traditional definitions of authorship and originality are grounded in human creativity. However, as AI becomes increasingly capable of autonomous creation, these definitions must evolve. Legal systems around the world are already grappling with this issue, with varying approaches. Some jurisdictions are considering AI as a tool under the control of a human author, while others are exploring more radical reconceptions of authorship and creativity.

As we stand at the intersection of artificial intelligence and copyright law, the journey forward requires not only adaptation but also proactive innovation. The rapid evolution of AI in creative processes presents both challenges and opportunities for the legal system to protect intellectual property while fostering an environment of innovation and creative freedom. The following recommendations offer concrete steps towards achieving a balanced and forward-looking copyright framework:

5.1 Establish Clear Guidelines for AI Authorship

Legal systems worldwide should consider establishing clear guidelines that define the criteria for authorship in the context of AI-generated works. This includes distinguishing between works where AI acts as a tool under human direction and works generated autonomously by AI. By doing so, copyright law can better address issues of originality and creativity, ensuring that human creators are duly recognized and protected.

5.2 Create a Special Category for AI-Generated Works

Consider the creation of a special category within copyright law for AI-generated works. This category could offer a modified form of protection that acknowledges the unique nature of these works, possibly involving a shorter term of copyright or specific rights tailored to encourage sharing and further creative use. This approach would recognize the contributions of AI to the creative process without undermining the incentives for human creators.

5.3 Encourage Transparency and Attribution

Encourage transparency in the use of AI in creative processes by mandating attribution to both the human creator and the AI system used. This would not only provide clarity on the origins of a work but also foster an environment where creators are informed about the contributions of AI to their creative endeavors.

5.4 Foster International Collaboration

Given the global nature of both AI technology and the creative industries, international collaboration is essential. Countries should work towards harmonizing copyright laws to address AI-generated content, ensuring that creators and innovators have a consistent legal framework that supports their work across borders.

5.5 Promote Ethical Use of AI in Creativity

Alongside legal adaptations, there should be a concerted effort to promote the ethical use of AI in creative processes. This includes ensuring that AI systems are trained on diverse and non-infringing datasets and that the use of AI respects the cultural and moral rights of human creators.

5.6 Engage in Ongoing Dialogue with Stakeholders

Finally, the legal community should engage in ongoing dialogue with technologists, creators, and policymakers to understand the evolving capabilities of AI and its impact on creativity. This dialogue should inform continuous updates to copyright law, ensuring it remains relevant and responsive to technological advancements.

In conclusion, the path forward for copyright law in the age of artificial intelligence is one of careful consideration and collaborative effort. By taking tangible steps to adapt legal frameworks, we can ensure that copyright law continues to fulfill its fundamental purpose of promoting creativity and innovation, respecting the contributions of both human and artificial creators. As AI continues to shape the landscape of creative expression, our legal systems must evolve in tandem, offering clear, fair, and effective protection for the next generation of creative works.