Prediction of Mortality and Intervention in COVID-19 Patients Using Generative Adversarial Networks

Uiwon Hwang [1] **Euideuk Hwang** [2] **Minsoo Kang** [1] **Sungroh Yoon** [1]

Abstract

The COVID-19 pandemic hits worldwide with a significant number of deaths and poses a major threat to public health. Accurate predictions of the risk of death and medical interventions are crucial for the survival of infected patients and the distribution of limited medical resources. Although machine learning classifiers can be used to predict mortality and medical interventions, it is problematic to employ the methods because training data are limited whose attributes may be missing and classes may be imbalanced. To effectively cope with these problems, we construct HexaGAN with a hint mechanism to predict the survival of the patients and medical interventions such as intubation and supplemental oxygen. In experiments, our method outperforms combinations of existing techniques for limited data problems. Notably, our method showed about twice higher performance than benchmarks in predicting deceased patients correctly. We anticipate that our approach could help provide appropriate treatments on time, allocate limited medical resources efficiently, and ultimately reduce the mortality rate of COVID-19 patients.

1. Introduction

The coronavirus disease 2019 (COVID-19) has spread around the world since December 2019. The explosion in the number of patients causes a shortage of medical resources including medical staffs and hospital beds (Arabi et al., 2020), so accurate screening of patients who are at high risk of death or who require medical interventions helps efficient allocation of medical resources. In addition, timely clinical interventions, such as intubation and supplemental oxygen, are important to reduce inpatient mortality.

However, it is difficult for medical staffs to make decisions about treatment in real time and to allocate medical resources efficiently due to the large number of COVID-19 patients. This has increased the demand for automation of the decision-making process using patients' electronic health records (EHR). To meet the demand, automated tools, especially deep learning (DL) methods, to predict the risk of death and interventions have been actively proposed (Sankaranarayanan et al., 2021; Banoei et al., 2021; Varzaneh et al., 2022).

Although DL models for classification and prediction have remarkably advanced (Huang et al., 2017; Tan & Le, 2019), limited data problems in EHRs have stunted the application of the DL-based methods. There are three typical limited data problems in EHRs. First, some attributes can be missing (missing data). Second, outcomes can be missing due to the labeling cost (missing label). Third, outcomes can be imbalanced (class imbalance). For these reasons, DL models for predicting various outcomes related to COVID-19 may not perform best.

In order to address these problems, machine learning techniques for imputation, oversampling, and semi-supervised learning can be applied. Recently, deep generative models for limited data problems have been proposed (Salimans et al., 2016; Li et al., 2017; Yoon et al., 2018). Hwang et al. (2019) suggested a unified view of limited data problems via imputation and proposed generative adversarial networks (GANs) that can address the three problems simultaneously.

In this paper, we used a publicly available dataset (Cohen et al., 2020) of chest X-ray images and metadata to predict mortality and interventions of COVID-19 patients. We adopted HexaGAN (Hwang et al., 2019) and additionally applied a hint mechanism (Yoon et al., 2018) to accurately predict the mortality, and whether the patients need intubation or supplemental oxygen. We verified that our method outperforms combinations of existing techniques for limited data problems.

We can summarize our contributions as follows:

- We construct deep generative models to provide ac-

[1]Department of Electrical and Computer Engineering, Seoul National University, Seoul, Korea [2]School of Medicine, Kyungpook National University, Daegu, Korea. Correspondence to: Sungroh Yoon <sryoon@snu.ac.kr>.

Proceedings of the 1st Workshop on Healthcare AI and COVID-19, ICML 2022.

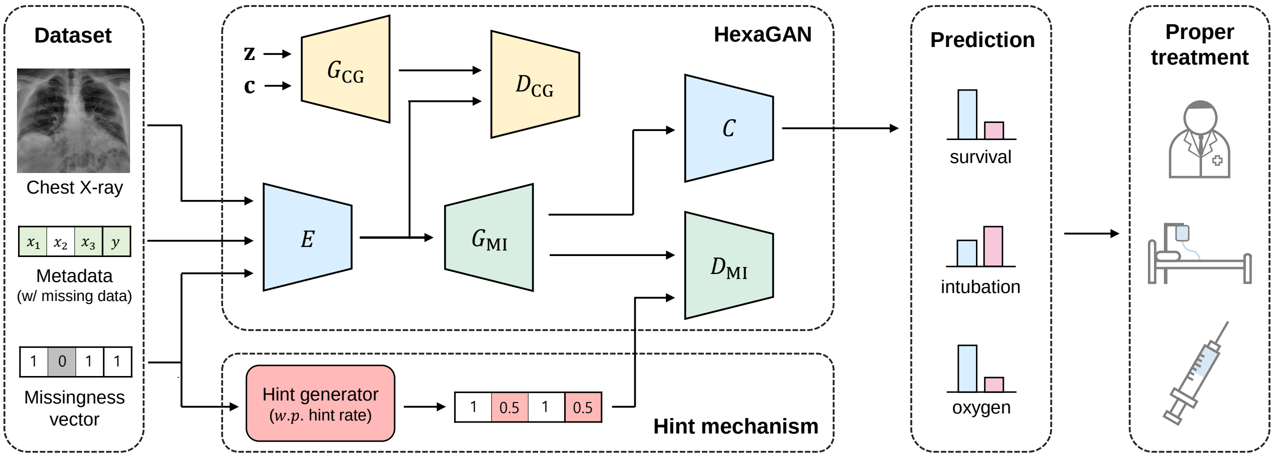

Figure 1. Overview of prediction of mortality and intervention in COVID-19 patients using GANs.

curate prediction of mortality and interventions for COVID-19 patients from the dataset with limited data problems.

- To enable this, we use generative adversarial networks named HexaGAN which addresses limited data problems simultaneously and additionally apply a hint mechanism to enhance the prediction performance.
- Our method significantly outperforms combinations of existing methods. Especially, our method achieves about twice higher performance (specificity) than benchmark combinations in predicting the risk of death.

2. Background

2.1. Coronavirus disease 2019

The coronavirus disease 2019 (COVID-19) pandemic poses an unprecedented threat to health and public health around the world. Since the outbreak in December 2019, the number of confirmed cases has exceeded 500 million, and the death toll from COVID-19 has surpassed 6 million people (up to May 16, 2022) (Dong et al., 2020; Wynants et al., 2020). Despite policies aimed at curbing the spread of the disease, numerous countries have faced medical crises, and there has been a marked shortage of medical resources, including beds and medical staff (Arabi et al., 2020). In fact, about 10% of SARS-CoV-2 infections are silent, the remaining 40% cause benign upper respiratory diseases, and about 20% cause pneumonia (Sah et al., 2021; Oran & Topol, 2020). About 10% of cases show hypoxemic pneumonia, which generally requires hospitalization for oxygen treatment (Zhang et al., 2022). In approximately 3% of cases, high-flow oxygen (O2 > 6 min^{-1}), mechanical ventilation (noninvasive or intubation) or extracorporeal membrane oxygenation (ECMO) is required (Zhang et al., 2020; Bastard et al., 2020). The provision of intervention in the intensive care unit is limited due to the shortage of medical resources, and the number of patients in need of intervention is explosively increasing, placing a burden on the medical system (Pei et al., 2021).

A remarkable epidemiological feature of COVID-19 is its strong age dependence, with a 10,000-fold greater risk in people over 80 years of age compared to people under the age of 10 (Zhang et al., 2022). The mortality rate from infection increases with age, from 0.001% for people aged 5-9 years to 8.29% for people aged 80 or older (O'Driscoll et al., 2021). However, age is not the only factor influencing mortality, and it is impossible to accurately predict patient mortality based on age information alone (Dessie & Zewotir, 2021; Takahashi et al., 2020; Hashim et al., 2020). Information on the prognosis of a disease is necessary to efficiently allocate medical resources and provide the best treatment to patients. In addition, timely clinical interventions such as intubation and supplemental oxygen are important to decrease the death rate of inpatients (Vera et al., 2021; Long et al., 2021). Therefore, machine learning models that predict the need for intervention and prognosis can be of great help in responding to public health crises and saving patients' lives. The interest in dealing with COVID-19 patient data is increasing, and among them, methods using deep learning are being actively proposed (Sankaranarayanan et al., 2021; Banoei et al., 2021; Varzaneh et al., 2022).

The most significant complication of severe COVID-19 is the acute respiratory distress syndrome (ARDS), which is associated with a high mortality rate of 35%-46% (Thompson et al., 2017; Annane et al., 2017; FakhriRavari et al., 2021). The COVID-19 treatment guidelines of National Institutes of Health (NIH) provide treatment guidance for adults hospitalized for COVID-19 based on their disease severity and oxygen requirements. In the case of severe

Table 1. Description of metadata in the dataset

Attribute	Description	Data type
age	Age	integer
offset	Days elapsed since the onset of symptoms or hospitalization	integer
sex	Male or female	binary
RT-PCR positive	Yes or no	binary
went icu	Whether the patient was in CCU (critical care unit) or ICU (intensive care unit)	binary
extubated	Whether the patient was successfully extubated	binary
temperature	Temperature of the patient ($^{\circ}$C)	continuous
pO2 saturation	Partial pressure of oxygen saturation (%)	continuous
wbe count	White blood cell count (10^3/uL)	continuous
neutrophil count	Neutrophil cell count (10^3/uL)	continuous
lymphocyte count	Lymphocyte cell count (10^3/uL)	continuous
view	Anteroposterior (AP), Posteroanterior (PA), Lateral (L) or AP Supine (APS) for X-rays	categorical
survival	Yes or no	binary
intubated	Whether the patient was intubated (or ventilated)	binary
supplemental O2	Whether the patient required supplemental oxygen	binary

COVID-19 patients, antiviral and immunomodulatory therapy using remdesivir and dexamethasone, and anticoagulant treatment using prophylactic dose of heparin are recommended. In particular, high-dose dexamethasone may be effective in patients developing ARDS (NIH, 2022). The use of dexamethasone is preferred for COVID-19 patients who are hospitalized and need supplemental oxygen, mechanical ventilation, or ECMO (Group, 2021; FakhriRavari et al., 2021). Indeed, many studies suggest that doctors should not use corticosteroids in early disease treatment processes when patients do not require oxygen support due to potential harms, while others have shown that early administration of dexamethasone prevents the progression to a severe disease, without increased mortality (FakhriRavari et al., 2021; Lee et al., 2021; Arora & Panda, 2021).

2.2. Learning from limited data

Machine learning (ML) classifiers have shown outstanding performances when trained on clean data (Huang et al., 2017; Tan & Le, 2019). However, the data collected in real-world can be limited, and it hinders the performance of the classifiers. There are three limited data problems in EHRs, and ML-based preprocessing techniques have been utilized to solve these problems. First, records where some features are missing cannot be used directly for learning a model. One strategy to handle such missing features is to fill them with certain values before training. Several well-known imputation techniques are mean imputation, k-nearest neighbors (kNN) (Troyanskaya et al., 2001), multivariate imputation by chained equations (MICE) (Van Buuren & Groothuis-Oudshoorn, 2011). Second, class labels in training data can be missing, and semi-supervised learning is performed to utilize data without labels to train the model. Semi-supervised learning methods include kNN and label propagation (LP) (Zhu & Ghahramaniн, 2002). Third, when the classes of training data are imbalanced, the predictions of the trained model can be biased towards the majority class. Oversampling techniques, including the synthetic minority oversampling technique (SMOTE) (Chawla et al., 2002) and adaptive synthetic (ADASYN) (He et al., 2008), and regularization methods, including the cost sensitive (CS) loss (Sun et al., 2007) and the class rectification loss (CRL) (Dong et al., 2017), have been proposed to solve the class imbalance problem. These methods can only deal with one of the limited data problems. When several limited data problems exist at the same time in a training dataset, a cascade combination of these methods can be used.

2.3. Generative adversarial networks

Generative adversarial networks (GANs) are one of deep generative models which implicitly estimate data distribution via an adversarial learning between a generator and a discriminator (Goodfellow et al., 2014). Typically, GANs use the adversarial losses as follows:

$$\mathcal{L}_D = -\mathbb{E}_{p(\mathbf{x})}\left[f(D(\mathbf{x}))\right] - \mathbb{E}_q\left[g(D(G(\mathbf{z})))\right] \quad (1)$$

$$\mathcal{L}_G = \mathbb{E}_{q(\mathbf{x})}\left[h(D(G(\mathbf{z})))\right] \quad (2)$$

where $f, g, h : \mathbb{R} \mapsto \mathbb{R}$ are loss metrics. For example, $f(t) = \log\sigma(t)$, $g(t) = h(t) = \log(1-\sigma(t))$ for vanilla GANs (Goodfellow et al., 2014) and $f(t) = t$, $g(t) = h(t) = -t$ for Wasserstein GANs (Arjovsky et al., 2017). Here, $p(\mathbf{x})$ denotes the real data distribution, $q(\mathbf{x})$ denotes the generated data distribution, and σ is the sigmoid function. The generator and discriminator are learned by minimizing $\mathcal{L}_G$ and $\mathcal{L}_D$, respectively. GANs have been mainly developed as models for image generation (Karras et al., 2019) and speech synthesis (Kong et al., 2020). They generate realistic data to the extent that humans cannot distinguish them from real data.

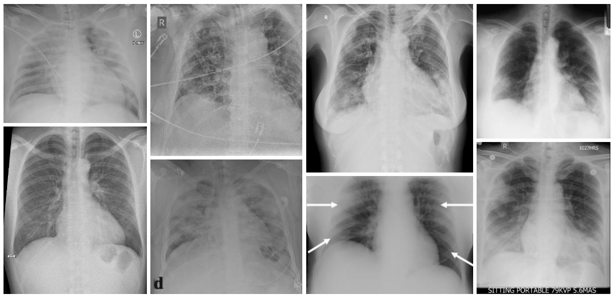

Figure 2. Samples of X-ray image data of COVID-19 patients.

Taking advantage of the excellent generation capability of GANs, attempts have been made to solve the limited data problems. GAN-based methods for missing data imputation (Yoon et al., 2018), oversampling (Engelmann & Lessmann, 2021), and semi-supervised learning (Li et al., 2017) have been proposed. Among them, GAIN (Yoon et al., 2018) is a method for missing data imputation. The original GAN discriminator distinguishes between real and generated records, whereas the GAIN discriminator distinguishes between missing and non-missing elements in element-wise manner. In addition, the imputation performance was improved by introducing a hint mechanism. As one of the most notable study, HexaGAN (Hwang et al., 2019) considers a situation where three limited data problems co-exist in the training data at the same time. HexaGAN defines limited data problems in terms of imputation and addresses the problems simultaneously.

3. Method

We used the HexaGAN framework (Section 3.1) to address limited data problems and applied a hint mechanism (Section 3.2) to enhance the prediction performance. The overview of our proposed method is depicted in Figure 1.

3.1. HexaGAN

For our task, the dataset contains chest X-ray image data $\mathbf{I}$, tabular metadata $\mathbf{t}$, and outcomes $\mathbf{y}$. We construct a boolean vector $\mathbf{m}$ named a missingness vector to indicate whether an element of metadata is missing. If i-th element of metadata is missing, m_i is 0, and vice versa. Since the dataset has the missing data, class imbalance, and missing label problems simultaneously, we adopted the HexaGAN framework (Hwang et al., 2019) which is the state of the art method for this scenario.

HexaGAN consists of six components including an encoder, generators, discriminators, and a classifier. These components interact with each other via adversarial learning and low-dimensional vectors $\mathbf{h}$ in the hidden space. Role of each component is summarized as follows:

Table 2. The number of outcomes with rates of missing data

Outcome	Y	N	Missing	Imb. ratio (1:x)	Missing rate
survival	162	38	384	4.3	65.8%
intubated	114	76	394	1.5	67.5%
supplemental O2	53	20	511	2.7	87.5%

- The encoder E receives $\mathbf{I}$, $\mathbf{t}$, and $\mathbf{m}$ and synthesizes a low-dimensional hidden vector $\mathbf{h}$ in the hidden space.
- A generator for missing imputation G_{MI} receives $\mathbf{h}$ and performs missing data imputation and oversampling.
- A discriminator for missing imputation D_{MI} receives imputed data and predicts $\mathbf{m}$.
- A generator for conditional generation G_{CG} receives a minority class label $\mathbf{c}$ and a noise vector $\mathbf{z}$ and generates $\mathbf{h}$ to oversample data in the minority class.
- A discriminator for conditional generation D_{CG} distinguishes between $\mathbf{h}$ from real data and $\mathbf{h}$ synthesized by G_{CG}.
- The classifier C provides predictions given imputed data and generates pseudo-label of unlabeled data to perform semi-supervised learning.

As described in Hwang et al. (2019), these components interact with each other to address the limited data problems. Each component of the whole framework is updated in rotation (See Algorithm 2 in Hwang et al. (2019)).

3.2. Hint mechanism

A hint mechanism, which is proposed by Yoon et al. (2018), is a random variable containing partial information about missingness. We used the hint mechanism to improve the imputation performance of our method, which eventually enhanced the prediction performance of the classifier. When the hint rate is p which is the hyperparameter, each element of a hint mechanism $\mathbf{H}$ is sampled as follows:

$$H_i = \begin{cases} m_i & w.p.\ p \\ 0.5 & w.p.\ 1-p \end{cases} \tag{3}$$

A sampled $\mathbf{H}$ is fed into D_{MI}. With little computational overhead, it helps D_{MI} predict $\mathbf{m}$, and improve the imputation performance of our method.

4. Experiments

We used the covid-chestxray-dataset (Cohen et al., 2020) for the prediction of mortality and medical interventions. The dataset includes 846 records who have been confirmed or

Table 3. Mortality prediction performance of combinations of existing techniques

Missing data	Class imbalance	Missing label	F1-score	Accuracy	AUROC	Sensitivity	Specificity
Mean	SMOTE	kNN	0.9162	0.8570	0.8490	0.9748	0.3909
kNN	SMOTE	kNN	**0.9411**	**0.8948**	**0.8893**	0.9861	0.3748
MICE	SMOTE	kNN	0.9198	0.8535	0.8093	0.9972	0.0826
Mean	ADASYN	kNN	0.8992	0.8251	0.8266	0.9778	0.2232
kNN	ADASYN	kNN	0.9367	0.8854	0.8888	0.9945	0.2631
MICE	ADASYN	kNN	0.9159	0.8464	0.7736	0.9930	0.0595
Mean	CRL	kNN	0.8961	0.8180	0.7888	0.9837	0.1642
kNN	CRL	kNN	0.9333	0.8794	0.8067	0.9847	0.2729
MICE	CRL	kNN	0.9181	0.8522	0.7612	0.9818	0.1567
Mean	CS	kNN	0.9029	0.8321	0.7860	0.9763	0.2634
kNN	CS	kNN	0.9213	0.8546	0.7193	**1.0000**	0.0240
MICE	CS	kNN	0.9162	0.8475	0.7100	0.9874	0.0963
Mean	SMOTE	LP	0.8713	0.7778	0.7781	0.9815	0.1119
kNN	SMOTE	LP	0.8762	0.7920	0.8163	0.9905	0.2201
MICE	SMOTE	LP	0.8839	0.8085	0.8307	0.9610	0.3317
Mean	ADASYN	LP	0.9015	0.8381	0.8036	0.9676	**0.4141**
kNN	ADASYN	LP	0.8739	0.7932	0.8132	0.9666	0.2926
MICE	ADASYN	LP	0.8746	0.7836	0.7914	0.9953	0.1220
Mean	CRL	LP	0.8740	0.7790	0.7525	**1.0000**	0.0558
kNN	CRL	LP	0.8744	0.7943	0.7886	0.9569	0.3231
MICE	CRL	LP	0.8739	0.7943	0.8164	0.9392	0.3415
Mean	CS	LP	0.8906	0.8144	0.7579	0.9815	0.2683
kNN	CS	LP	0.8718	0.7825	0.7625	0.9921	0.1793
MICE	CS	LP	0.8751	0.7896	0.8287	0.9719	0.2195

suspected of COVID-19 or other viral and bacterial pneumonias. Among them, 468 records are obtained from COVID-19 patients, and those with outcome attributes were used as labeled data. Records obtained from other pneumonias patients were treated as unlabeled data regardless of the missingness of outcome attributes. We used X-ray image data and metadata which contain the patients' information. The descriptions of metadata attributes are shown in Table 1. We used the bottom three attributes as outcomes. For example, we used *survival* to predict mortality, *intubated* to predict intubation, and *supplemental O2* to predict supplemental oxygen. The samples of X-ray image data are shown in Figure 2, and the number of outcomes and missing rates are presented in Table 2.

We computed the F1-score, the area under the receiver operating characteristic curve (AUROC), the sensitivity, and the specificity as the evaluation metrics. We used 5-fold cross validation and reported the average of the test performance for each fold.

4.1. Mortality prediction

Benchmark combinations To find the most competent combinations of existing techniques for limited data problems, we conducted extensive experiments on mortality prediction as shown in Table 3. We used mean imputation (Mean), k-nearest neighbors (kNN) (Troyanskaya et al., 2001), and multivariate imputation by chained equations (MICE) (Van Buuren & Groothuis-Oudshoorn, 2011) for the missing data problem; synthetic minority oversampling technique (SMOTE) (Chawla et al., 2002), adaptive synthetic (ADASYN) (He et al., 2008), class rectification loss (CRL) (Dong et al., 2017), and cost sensitive loss (CS) (Sun et al., 2007) for the class imbalance problem; kNN and label propagation (LP) (Zhu & Ghahramaniн, 2002) for the missing label problem. As a result, a combination of kNN for the missing data problem, SMOTE for the class imbalance problem, and kNN for the missing label problem shows the highest performance in terms of the F1-score, accuracy, and AUROC. A combination of Mean for the missing data problem, ADASYN for the class imbalance problem, and LP for the missing label problem shows the highest performance in terms of specificity. Specificity is an important performance metric for mortality prediction. Since the outcome is 1 when the patient is alive and the outcome is 0 when the patient is dead, the specificity becomes the probability of correctly identifying patients who have died, which is an important indicator for accurately predicting patients who need urgent treatment. Therefore, we chose these two combinations as competent benchmark combinations to compare with our method.

Effectiveness of hint Figure 3 shows the benefit of a hint mechanism on mortality prediction. We evaluated the perfor-

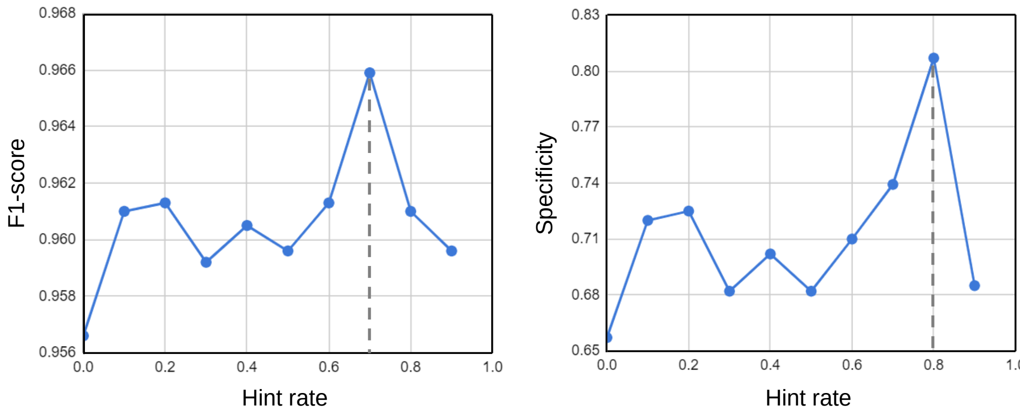

Figure 3. Performance comparison with respect to the hint rate.

Table 4. Performance comparison on mortality prediction

Method	Hint rate	F1-score	AUROC	Sensitivity	Specificity
kNN + SMOTE + kNN	N/A	0.9411	0.8893	0.9861	0.3748
Mean + ADASYN + LP	N/A	0.9015	0.8036	0.9676	0.4141
HexaGAN	0	0.9566	0.9395	**1.0000**	0.6571
HexaGAN w/ hint	0.7	**0.9659**	**0.9508**	**1.0000**	0.7393
HexaGAN w/ hint	0.8	0.9610	0.9464	0.9714	**0.8071**

mances of our method by changing hint rate from 0 to 0.9. If the hint rate is 0, it is the same as HexaGAN without a hint mechanism. The hint mechanism improved the prediction performance of HexaGAN. F1-score is best at hint rate 0.7, and specificity is best at hint rate 0.8. Therefore, we chose these two hint rates to conduct comparative studies.

Comparison on mortality prediction Table 4 compares the performance of our method with benchmarks on mortality prediction. Our method (HexaGAN with a hint mechanism) outperformed benchmark combinations. Especially, our method with a hint rate 0.8 performed approximately twice better than benchmark combinations in terms of specificity, indicating that it predicted critically ill patients approximately two times better.

4.2. Intervention prediction

We also predicted medical interventions including intubation and supplemental oxygen using our method, and conducted comparative studies between our method and benchmark combinations.

Comparison on intubation prediction Table 5 compares the performance of our method with benchmark combinations on intubation prediction. We did not use the *extubated* attribute as a predictor for intubation prediction, because a patient can only be extubated if they were intubated at some point. The combination of Mean + ADASYN + LP failed to predict intubation of patients. Our method showed the best performance across all evaluation metrics. HexaGAN with hint rates 0, 0.7, 0.8 showed the sensitivity of 1.0000, which indicates that our method predicts all patients who are actually intubated as patients in need of intubation.

Comparison on supplemental O2 prediction Table 6 compares the performance of our method with benchmarks on supplemental oxygen prediction. The combination of Mean + ADASYN + LP showed the highest performance in terms of F1-score and our method showed the best performance in terms of all evaluation metrics except for F1-score. HexaGAN with hint rates 0, 0.7, 0.8 showed the AUROC of 1.0000, which indicates that our method can perfectly predict supplemental oxygen by adjusting the prediction threshold.

5. Discussion

5.1. Machine learning perspective

In order to learn a classifier from the training data in which the limited data problems exist simultaneously, preprocessing techniques to solve each problem must be applied sequentially. However, this approach requires building a preprocessing pipeline that depends on problems present in the

Table 5. Performance comparison on intubation prediction

Method	Hint rate	F1-score	AUROC	Sensitivity	Specificity
kNN + SMOTE + kNN	N/A	0.9513	0.9490	0.9805	0.7826
Mean + ADASYN + LP	N/A	0.7133	0.6115	0.9742	0.0841
HexaGAN	0	0.9563	0.9729	**1.0000**	0.8227
HexaGAN w/ hint	0.7	0.9604	**0.9771**	**1.0000**	0.8409
HexaGAN w/ hint	0.8	**0.9648**	0.9764	**1.0000**	**0.8591**

Table 6. Performance comparison on supplemental O2 prediction

Method	Hint rate	F1-score	AUROC	Sensitivity	Specificity
kNN + SMOTE + kNN	N/A	0.9537	0.9771	0.9459	0.7493
Mean + ADASYN + LP	N/A	**0.9730**	0.9713	0.9744	0.8000
HexaGAN	0	0.9513	**1.0000**	**1.0000**	0.7500
HexaGAN w/ hint	0.7	0.9609	**1.0000**	**1.0000**	0.8000
HexaGAN w/ hint	0.8	0.9568	**1.0000**	0.9556	**0.9000**

training data. Because each technique focuses on a problem, it ignores the connections between the problems. However, HexaGAN has the advantage that there is no need to build a specific preprocessing pipeline. In addition, HexaGAN deals with limited data problems simultaneously through interaction between its components, which helps to maximize the prediction performance.

As mentioned in Hwang et al. (2019), the limited data problems can be solved via imputation. Existing ML methods to overcome the limited data problems either use the value of the nearby training data or perform only simple processing. This may result in the imputed data not faithfully following the data distribution. On the other hand, deep learning methods can learn complex and nonlinear patterns inherent in data. Our method estimates the data distribution using deep generative models and addresses the limited data problems more effectively. This eventually improves the prediction performance of mortality and interventions in COVID-19 patients.

5.2. Medical perspective

The model presented in this study can predict survival, the use of supplemental oxygen, and intubation which are closely related to the severity of COVID-19. Since the use of dexamethasone is important for the development and mortality of the disease (NIH, 2022), the effect of early use of dexamethasone according to the prediction of the model could be further investigated.

Since the dataset used for experiments was obtained from a cross-sectional study, the learned model can predict whether an intervention was observed in retrospective data. We expect that GAN methods trained on datasets obtained from longitudinal studies can predict whether an intervention should be recommended at a particular moment. This will be an important avenue for future work to provide more timely treatment to COVID-19 patients in the presence of the limited data problems.

6. Conclusion

We have proposed a method using HexaGAN and a hint mechanism to predict mortality and the need for medical interventions. Our method robustly predicts mortality and interventions in the presence of the limited data problems such as missing data, missing label, and class imbalance problems. We also confirmed that a hint mechanism can enhance the prediction performance of HexaGAN. We believe that the excellent performance on mortality and intervention prediction of our model can help allocate limited medical resources efficiently, provide timely and appropriate clinical interventions, and ultimately save more lives from COVID-19.

Acknowledgements

We would like to thank the anonymous reviewers of this manuscript and the members of Data Science and Artificial Intelligence Lab at SNU for their helpful discussions in our manuscript preparation.

References

Annane, D., Pastores, S. M., Rochwerg, B., Arlt, W., Balk, R. A., Beishuizen, A., Briegel, J., Carcillo, J., Christ-Crain, M., Cooper, M. S., et al. Guidelines for the diagnosis and management of critical illness-related corticosteroid insufficiency (circi) in critically ill patients (part i): Society of critical care medicine (sccm) and european society of intensive care medicine (esicm) 2017. *Intensive care medicine*, 43(12):1751–1763, 2017.

Arabi, Y. M., Murthy, S., and Webb, S. Covid-19: a novel coronavirus and a novel challenge for critical care. *Intensive care medicine*, 46(5):833–836, 2020.

Arjovsky, M., Chintala, S., and Bottou, L. Wasserstein generative adversarial networks. In *International conference on machine learning*, pp. 214–223. PMLR, 2017.

Arora, K. and Panda, P. K. Steroid harms if given early in covid-19 viraemia. *BMJ Case Reports CP*, 14(2): e241105, 2021.

Banoei, M. M., Dinparastisaleh, R., Zadeh, A. V., and Mirsaeidi, M. Machine-learning-based covid-19 mortality prediction model and identification of patients at low and high risk of dying. *Critical Care*, 25(1):1–14, 2021.

Bastard, P., Rosen, L. B., Zhang, Q., Michailidis, E., Hoffmann, H.-H., Zhang, Y., Dorgham, K., Philippot, Q., Rosain, J., Béziat, V., et al. Autoantibodies against type i ifns in patients with life-threatening covid-19. *Science*, 370(6515):eabd4585, 2020.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.

Cohen, J. P., Morrison, P., and Dao, L. Covid-19 image data collection. *arXiv 2003.11597*, 2020. URL `https://github.com/ieee8023/covid-chestxray-dataset`.

Dessie, Z. G. and Zewotir, T. Mortality-related risk factors of covid-19: a systematic review and meta-analysis of 42 studies and 423,117 patients. *BMC Infectious Diseases*, 21(1):1–28, 2021.

Dong, E., Du, H., and Gardner, L. An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases*, 20(5):533–534, 2020.

Dong, Q., Gong, S., and Zhu, X. Class rectification hard mining for imbalanced deep learning. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1851–1860, 2017.

Engelmann, J. and Lessmann, S. Conditional wasserstein gan-based oversampling of tabular data for imbalanced learning. *Expert Systems with Applications*, 174:114582, 2021.

FakhriRavari, A., Jin, S., Kachouei, F. H., Le, D., and Lopez, M. Systemic corticosteroids for management of covid-19: Saving lives or causing harm? *International Journal of Immunopathology and Pharmacology*, 35:20587384211063976, 2021.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

Group, R. C. Dexamethasone in hospitalized patients with covid-19. *New England Journal of Medicine*, 384(8): 693–704, 2021.

Hashim, M. J., Alsuwaidi, A. R., and Khan, G. Population risk factors for covid-19 mortality in 93 countries. *Journal of epidemiology and global health*, 10(3):204, 2020.

He, H., Bai, Y., Garcia, E. A., and Li, S. Adasyn: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*, pp. 1322–1328. IEEE, 2008.

Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.

Hwang, U., Jung, D., and Yoon, S. Hexagan: Generative adversarial nets for real world classification. In *International Conference on Machine Learning*, pp. 2921–2930. PMLR, 2019.

Karras, T., Laine, S., and Aila, T. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4401–4410, 2019.

Kong, J., Kim, J., and Bae, J. Hifi-gan: Generative adversarial networks for efficient and high fidelity speech synthesis. *Advances in Neural Information Processing Systems*, 33:17022–17033, 2020.

Lee, H. W., Park, J., Lee, J.-K., Park, T. Y., and Heo, E. Y. The effect of the timing of dexamethasone administration in patients with covid-19 pneumonia. *Tuberculosis and respiratory diseases*, 84(3):217, 2021.

Li, C., Xu, T., Zhu, J., and Zhang, B. Triple generative adversarial nets. *Advances in neural information processing systems*, 30, 2017.

Long, L., Wu, L., Chen, L., Zhou, D., Wu, H., Lu, D., Li, H., Duan, X., Han, Y., Li, X., et al. Effect of early oxygen therapy and antiviral treatment on disease progression in patients with covid-19: A retrospective study of medical charts in china. *PLoS neglected tropical diseases*, 15(1): e0009051, 2021.

NIH. Covid-19 treatment guidelines. https://www.covid19treatmentguidelines.nih.gov/management/clinical-management/clinical-management-summary/, 2022. [Online; accessed 31-May-2022].

Oran, D. P. and Topol, E. J. Prevalence of asymptomatic sars-cov-2 infection: a narrative review. *Annals of internal medicine*, 173(5):362–367, 2020.

O'Driscoll, M., Ribeiro Dos Santos, G., Wang, L., Cummings, D. A., Azman, A. S., Paireau, J., Fontanet, A., Cauchemez, S., and Salje, H. Age-specific mortality and immunity patterns of sars-cov-2. *Nature*, 590(7844):140–145, 2021.

Pei, S., Yamana, T. K., Kandula, S., Galanti, M., and Shaman, J. Burden and characteristics of covid-19 in the united states during 2020. *Nature*, 598(7880):338–341, 2021.

Sah, P., Fitzpatrick, M. C., Zimmer, C. F., Abdollahi, E., Juden-Kelly, L., Moghadas, S. M., Singer, B. H., and Galvani, A. P. Asymptomatic sars-cov-2 infection: A systematic review and meta-analysis. *Proceedings of the National Academy of Sciences*, 118(34), 2021.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. Improved techniques for training gans. *Advances in neural information processing systems*, 29, 2016.

Sankaranarayanan, S., Balan, J., Walsh, J. R., Wu, Y., Minnich, S., Piazza, A., Osborne, C., Oliver, G. R., Lesko, J., Bates, K. L., et al. Covid-19 mortality prediction from deep learning in a large multistate electronic health record and laboratory information system data set: Algorithm development and validation. *Journal of medical Internet research*, 23(9):e30157, 2021.

Sun, Y., Kamel, M. S., Wong, A. K., and Wang, Y. Cost-sensitive boosting for classification of imbalanced data. *Pattern recognition*, 40(12):3358–3378, 2007.

Takahashi, T., Ellingson, M. K., Wong, P., Israelow, B., Lucas, C., Klein, J., Silva, J., Mao, T., Oh, J. E., Tokuyama, M., et al. Sex differences in immune responses that underlie covid-19 disease outcomes. *Nature*, 588(7837): 315–320, 2020.

Tan, M. and Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pp. 6105–6114. PMLR, 2019.

Thompson, B. T., Chambers, R. C., and Liu, K. D. Acute respiratory distress syndrome. *New England Journal of Medicine*, 377(6):562–572, 2017.

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., and Altman, R. B. Missing value estimation methods for dna microarrays. *Bioinformatics*, 17(6):520–525, 2001.

Van Buuren, S. and Groothuis-Oudshoorn, K. mice: Multivariate imputation by chained equations in r. *Journal of statistical software*, 45:1–67, 2011.

Varzaneh, Z. A., Orooji, A., Erfannia, L., and Shanbehzadeh, M. A new covid-19 intubation prediction strategy using an intelligent feature selection and k-nn method. *Informatics in Medicine Unlocked*, 28:100825, 2022.

Vera, M., Kattan, E., Born, P., Rivas, E., Amthauer, M., Nesvadba, A., Lara, B., Rao, I., Espíndola, E., Rojas, L., et al. Intubation timing as determinant of outcome in patients with acute respiratory distress syndrome by sars-cov-2 infection. *Journal of critical care*, 65:164–169, 2021.

Wynants, L., Van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., Bonten, M. M., Dahly, D. L., Damen, J. A., Debray, T. P., et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *bmj*, 369, 2020.

Yoon, J., Jordon, J., and Schaar, M. Gain: Missing data imputation using generative adversarial nets. In *International conference on machine learning*, pp. 5689–5698. PMLR, 2018.

Zhang, Q., Bastard, P., Liu, Z., Le Pen, J., Moncada-Velez, M., Chen, J., Ogishi, M., Sabli, I. K., Hodeib, S., Korol, C., et al. Inborn errors of type i ifn immunity in patients with life-threatening covid-19. *Science*, 370(6515): eabd4570, 2020.

Zhang, Q., Bastard, P., Cobat, A., and Casanova, J.-L. Human genetic and immunological determinants of critical covid-19 pneumonia. *Nature*, 603(7902):587–598, 2022.

Zhu, X. and Ghahramaniн, Z. Learning from labeled and unlabeled data with label propagation. 2002.